

Understanding of Tourist Dynamics from Explicitly Disclosed Location Information

Fabien Girardin¹, Filippo Dal Fiore², Josep Blat¹, Carlo Ratti²

¹ Universitat Pompeu Fabra, Interactive Technologies Group
Passeig de Circumvalació 8, 08003 Barcelona, Spain
{Fabien.Girardin, Josep.Blat}@upf.edu

² Massachusetts Institute of Technology, SENSEable City Lab
77 Massachusetts Avenue, Cambridge, MA 02139 USA
{dalfiore, ratti}@mit.edu

Abstract. In recent years, the large deployment of mobile devices has led to a massive increase in the volume of records of where people have been and when they were there. The analysis of the accumulated archives of such spatio-temporal data can derive high-level human behavior information valuable to urban planners, traffic engineers, and tourism authorities. In this paper, we describe our approach to collect and analyze the history of physical presence of tourists from the digital footprints they publicly make available on the world-wide web. Our work takes context in the Province of Florence where besides data from survey-based hotel and museums frequentation, local authorities have limited information on the fluxes of visitors and on the nationalities of the tourists who do not sleep in town. As a proof of concept, we used a corpus of publicly disclosed geo-annotated photos taken in the province by 3348 photographers over a period of 2 years. Based on the time and location of the photos, we design geovisualizations to reveal the tourist activity and flows in space and time. Our initial results provide insights on the density of tourists in the area and well as on the flow and activity of tourists within, in and on of that area.

Keywords: Spatio-temporal data analysis, geovisualization, location-disclosure, location-based services.

1 Introduction

In the past years, research in location sensing and tracking has been dominated by figuring out where persons and objects are in space (see Hazas et al., (2004) for a review). The wide adoption of these mobile and wireless technologies often referred as Location-Based Services allow new means of supporting people perception of their surrounding physical and social spaces. In parallel, the records of where people have been and when they were (i.e. spatio-temporal data), produced by these services, have led to the understanding of different aspects of mobility and travel. The analysis of these data enable to recognize the modes of mobility (Sohn et al., 2006), infer travel purposes (Wolf, 2001), define significant places (Ashbrook and Starner, 2002), cluster tourist routes in a city (Asakura and Iryob, 2007), categorize urban areas from network activity patterns (Ratti et al, 2006) or predict a driver's destination as a trip progresses (Krumm and Horwitz, 2006).

Similar to these works, our approach benefits from people's experience of mobile devices, to gain a more thorough understanding of urban environments. However, we take advantage of the recent explosion in the use of capture devices (e.g. mobile phones, digital cameras) and collaborative web platforms to share their content (see Torniai et al., (2007) for a review). This user-generated information provide large amount of digital data linked to the physical world. Recent research works showed the potential of the geographically annotated material available on the Web. For instance, Ahern et al., (2007) and Snavely et al. (2006) developed means of world exploration via photos and maps (e.g. for "virtual tourism"). Using a similar dataset, Rattenbury et al. (2007) showed that the location and time metadata associated with photos and their tags enables the extraction of "place" and "event" semantics. In this study, we focus on the sense of presence based on a collection of photographs captured in trips. We consider that uploading, tagging and disclosing the location of a photo can be interpreted as an act of physical presence in

time (i.e. “I was here”). To prove the concept, we retrieved from the popular photo-sharing web platform Flickr¹ 85910 photos taken by 3348 photographers over a period of 2 years. Based on the time and the disclosed location of the photos, we extracted records of their presence; performed statistical analysis and designed geovisualizations. These outputs allow the evaluation of the potential of using people-generated geographically referenced information to contribute to understanding how people travel and people experiencing the city.

Inspired by Ratti et al. (2006), we perceive a few domains of applicability of the processing and visualization of these large personal logs. Firstly, they can provide urban planners, local authorities and traffic engineers with useful information on how a city gets used by different groups. Knowing who populates different parts of the city at different times can complement existing data sources on the basis of which decision and policy making normally takes place. Provision of customized services (or advertising), rescheduling of opening times or reallocation of existing service infrastructures, are examples of possible outputs. Visualizations can nonetheless be helpful to raise awareness among city users, as for the current ways in which they populate the city as well as for possible better alternatives.

To prove the applicability of the concept, we applied it in the context the Province of Florence, Italy. The local authorities desperately aim at better understanding the tourist fluxes inside its boundaries. So far, they have been using classic survey-based hotel and museums frequentation data to get to know where tourists of different nationalities prefer to spend their time (hence money) in Florence. However, they truly lack of observations on the mobility, nationality and quantity of the “day trippers”, that is the tourists who visit Florence but are "invisible" in the data, as they do not sleep in town.

In the reminder of this paper, we will first we acknowledge the current works in the mobility data collection and their constraints to perform travel surveys. These shortcomings suggest that our research community should investigate and evaluate new data and perspectives. In consequence, we propose a novel approach that takes advantage of spatio-temporal data generated by tourists when publicly sharing their photos on the world-wide web. Second, we describe the types of data that can be collected and their meanings. Third, we present the preliminary results of the analysis with the support of geospatial visualizations. Finally, we conclude with the assessment of this proof of concept and discuss its meaning for future works.

2 Related Works

The recent emergence of location technologies and techniques favored the development of new approaches to capture and analyze people’s mobility (see Wolf, 2004 for a survey). The aim has been to replace traditional travel diaries (e.g. paper-and-pencil interview, computer-assisted telephone interviews, computer-assisted-self-interview) by automatically collecting mobility data. For instance, Wolf et al. (2001) proved the feasibility of using Global Positioning System (GPS) data loggers to improve the quality or completely replace traditional surveys. However, this type of mobility survey faces the problem inherent to longitudinal studies such as recruiting a pool of respondents or preventing any fatigue effects. Besides the privacy concerns of continuously and precisely tracking people, Schoenfelder et al. (2002) and Stopher et al. (2003) identify the potential technical drawbacks of a GPS-based approach. They list transmission problems, warm-up times before getting a valid position, and the cost of post-processing of the GPS data as issues that impair the quality of the survey. Similarly, in their findings, Wolf and colleagues mention that the equipment packages deployed for their pilot study proved to have many more problems than anticipated. For instance, the off-the-shelf units and cabling used were not optimized for durability, resulting in losing participants due to problems with equipment performance and user operations.

Other mobility studies relied on the mobile phones use of the Global System for Mobile

¹ <http://www.flickr.com>

communications (GSM) network to generate mobility data. In a first type of approach, the mobile devices calculate and report their position. For instance, the TeleTravel System (TTS) project (Wermuth, 2001) combined a mobile device GSM tracking technology and an electronic travel diary to determine the travel behavior of the respondents. Similarly, Asakura and Iryo (2005) and Sohn et al. (2006) determined individuals travel behavior by interpreting changes in the set of nearby towers and signal strengths as indicative of position and motion. In a study of travel routines, Froehlich et al. (2006) augmented the GSM mobility data with context-triggered in-situ survey in which panelists would explicitly rate the place they are in. Another approach consists in measuring the traffic intensity and migration of each cell in a GSM network to capture the movement patterns of mobile phone users. Ratti et al. (2006) use this technique as a tool for urban planning in disclosing the evolution through space and time of the activities in a city. With comparable network data IntelliOne² provides TrafficAid a traffic mapping system developed that identifies congestion in real time. While this approach scales because it does not rely on a costly software or hardware deployment, it fails at collecting individual traces. Finally, recent studies have investigated the deployment of Bluetooth enabled systems to establish the flow of people at specific locations (O'Neill et al., 2006) as well as recognize daily user activities and identify socially significant locations (Eagle and Pentland, 2006).

However, there are several key issues with the use of technologies for collecting travel behavior; these include privacy concerns about collecting data without traveler consent, longevity (e.g. preventing fatigue effects), the ability to collect individual traces and the scalability challenge deploying the tracking system such as dealing with a variety of cellular network standards and providers. All the above-mentioned approaches in capturing mobility data face one or many of these aspects of performing travel surveys. Table 1 shows how GPS, GSM and Bluetooth tracking performs with important aspects of a travel survey (i.e. scalability, longitude, individual information, and privacy).

Table 1. Mobility data capture with the constraints of travel surveys

Mobility data capture	Scalability	Longitude	Individual	Privacy
GPS	Low	Low	High	High
GSM (device-based)	Low	Medium	High	High
GSM (aggregated network-based)	High	High	Low	Medium
Bluetooth	Low	High	High	Medium

Some scholars work in the combination of the approaches to produce more complete spatial data (origin, route, destination) when tracking individual's routes (Kracht, 2004). However, while promising, it has not proved to solve all the issues. In the reminder of this paper, we argue that a new type of explicitly disclosed spatio-temporal data coming from on public web platforms such as Flickr can overcome these four constraints and provide an additional set of spatio-temporal data to understand the dynamic of people in a urban space.

3 The data

We study the photos uploaded on to the popular photo-sharing platform Flickr. People use this service to share and organize photos with the option to geographical references them. Each time a photo is virtually linked to a physical location, the system assigns a longitude and latitude and retrieves the time of capture from the photo Exchangeable Image File Format (EXIF) metadata. The location generally marks where the photo was taken; but sometimes marks the area of the photographed object. An accuracy attribute referring to the zoom level of the map used to disclose the location of the photo (e.g. from 8 for region/city level to the most precise 16 for street

² <http://www.intellione.com/>

level) completes the spatio-temporal information. A chronologically ordered set of photos creates traces that reveal the movements of each individual in space. Figure 1 describes the process of recording and collecting the data. First, tourists take photos during their trips and journeys. Later, they manually associated a position to the photos through a map interface in Flickr or other external map-based services. A minority of tech savvy users has their photos implicitly annotated with a position extracted from data collected by GPS devices embedded or external to the camera. Finally, we retrieve via the Flickr API, the coordinates, timestamp, accuracy level, and an obfuscated identifier of the owner of the publicly available photos in a given area.

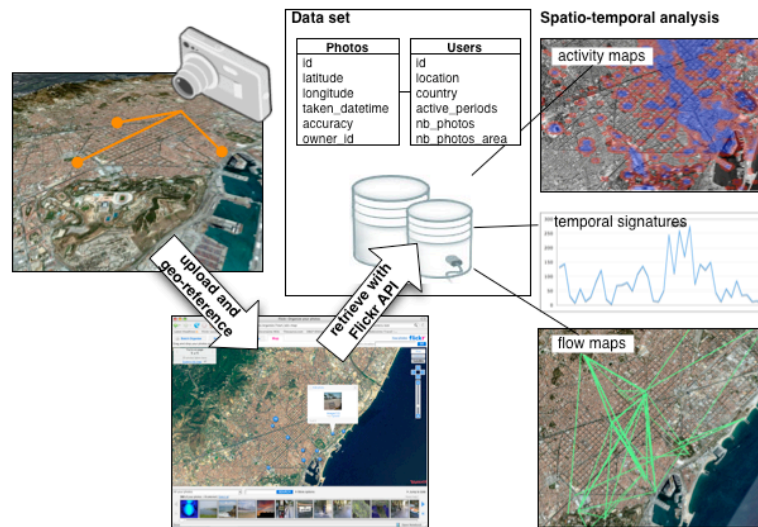


Figure 1. Data flow, from data recording, retrieving, storing to the visualizations. *Photographic imagery copyright 2007 DigitalGlobe.*

Our data set contains photos of the Province of Florence over a 2-years period (from April 2005 to April 2007). In a first step we had to separate visitors from inhabitants by determining the presence in time in the area as discriminating factor. We divided the time in arbitrary periods of 30 days and assigned the number of periods each photographer was active in the area. For example, if a photographer took all his or her photos within 30 days, the algorithm considers him or her as a visitor. On the other hand, when photos from the same photographer reveal an interval greater than 30 days, then he or she was categorized as inhabitant. The aim of this strict threshold was to capture the real on-time tourists. Out of a population of 3348 photographers, we identified 2880 one-time visitors.

In addition, we were interesting in knowing more about the nationalities of our panel of photographers. Hence, we took advantage of the Flickr social function that invites (but not obligates) users to provide “offline bits” on their city and country of residence. We found out that 65% (2146 out of 3348) of the users actually disclosed this information. While it is hard to predict how much of this data is truthful, in most cases the country of residence could be retrieved by programmatically parsing the data. In 11% (250/2146) of the cases, we had to manually assign a country, because of misspelling errors or eccentric name (e.g. “Big Apple” for New York).

Collecting and analyzing any kind of mobility data raise serious privacy issues; people are concerned about revealing the history of their whereabouts to un-trusted third party applications. Our approach assesses these utmost important uses at two different levels. First, the users explicitly disclose the position of their photos on a maps and control who gets access to their location data. Second, while obtaining this public information, we applied an obfuscation algorithm to lose the relationship with the identity used in Flickr. In consequence, we exclusively treat anonymized records of digital traces publicly disclosed by individuals.

Our experience of collecting Flickr geo-referenced data provided some insights on the data at hand. First, Flickr contains more than 20 million photos linked to a physical location (as of March 2007). Cities such as London or New York contain each more than 250'000 photos and counting (e.g. a growing pace of around 400 photos per day in London, around 150 per day in Barcelona). The quantitative richness of the Flickr data set presages the augmentation of publicly accessible people-generated location and time sensitive data. Second, a very specific type of people use Flickr to geographically reference their images. They are normally well traveled and technologically savvy. Therefore, we are dealing with a very specific type of tourists. Third, the timestamps extracted from the camera-generated EXIF metadata do not necessarily match the correct time. Indeed, the capturing devices might not be set with the correct time zone information. In consequence, there is high uncertainty on the accuracy of the time embedded in the photo when it is taken. From the temporal analysis we describe later, the date embedded in a photo seems to be less an issue. For example when uniquely considering the photos taken in 2006, we can assume that the user of the camera had to set the date and it was not the default value of the camera which is rarely assigned to 2006.

4 The results

This section of the paper presents the geovisualizations that are being produced and discusses their meanings. In the research process, we were driven by an overarching research question: what is the spatio-temporal behavior of tourists from different nationalities, inside the Province of Florence? Flickr was deemed an adequate context to extract data from, for it contains the spatial element, as well as the temporal and the nationality. But with the following caveats: firstly, there is not guarantee on their degree of accuracy; secondly we can not assume perfect mirroring between the revealed spatio-temporal patterns and the actual ones, for we can not know where users have actually been between a given photo and the next one. In other words, we can only deal with approximations. With these limitations in consideration, we cover an analysis of the data based on map visualizations to reveal the tourist activity and flow within the central Italy region and the city of Florence. Each set of maps aims at giving indications on the applicability of the collected data set to better understand the tourist dynamics within a certain area. First, two analysis aim at brining a new perspective on the spatial and temporal density of visitors in the Province of Florence and City of Florence: (a) Characterize the areas of the city/region: where are the tourists concentrated? (b) Reveal temporal signatures: activity by day-of-week and months of the year, and days of the year. The two latter geovisualizations seek a better understanding the flow of tourists into, out of and within the Province of Florence (c) Define patterns of flow in and out of the studies area, (d) Define the patterns of flow within the Province of Florence.

4.1 Where are the tourists concentrated?

To understand where tourist concentrate, we extracted from our dataset the latitude and longitude of all the images taken by visitors and geotagged with an accuracy of greater than 10 (i.e. city level). From this data, we produced a matrix with the number of photos taken in each of the area represented by each cell. To reveal the tourist concentrations, developed a solution performing a level plot with the R statistical software. In addition, we produced an interactive interface revealing an activity “heat map” on top of the Google Maps system (Figure 2). In this solution, we generated a Keyhole Markup Language (KML) document from the activity matrix and fed it to the Geocommons³ software. The results reveal a zoomable map of the overall tourist activity covering the northern part of central Italy (Figure 2a), including a part of the Tuscany region (Figure 2b), the city of Florence (Figure 2c) and around the famous Ponte Vecchio (Figure 2d). Currently, these geovisualizations give an overview of the activity. However, they do not offer a scale to understand the quantitative meaning of the colors.

³ <http://www.geocommons.com/>

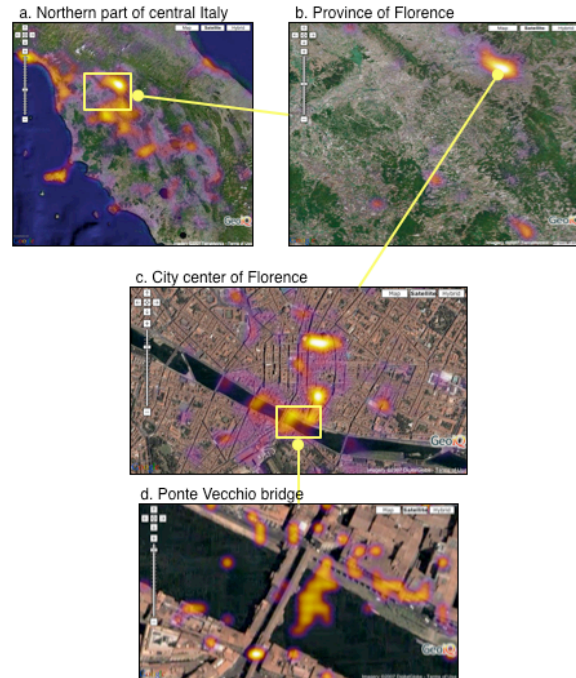


Figure 2. Interactive zoomable map of the overall tourist activity from (a) the northern part of central Italy to (d) fine-grained information on the Ponte Vecchio in Florence. *Photographic imagery copyright 2007 Cnes/Spot, NASA, DigitalGlobe and TerraMetrics.*

These visualizations show that the main activity in the Province of Florence actually takes place in the city of Florence. There are other major points of interest around the province, such as the cities of Pisa and Siena, as well as the Mediterranean coast between La Spezia and Livorno and the Island of Elba. Within the city, the tourists concentrate around the Ponte Vecchio and the Basilica di Santa Maria del Fiore.

4.2 Spatio-temporal signatures

In the first analysis, we produced activity maps regardless to the period when the analyzed data was recorded. In a next step, we generated similar maps to compare the activity in time such as the seasons and the weekends versus weekdays. Prior to develop spatio-temporal signature of the area of study, we took its “tourist pulse”. We charted the temporal signature of the number of visitors active in the. Figure 4 shows the weekly, monthly yearly pulse of the year 2006.

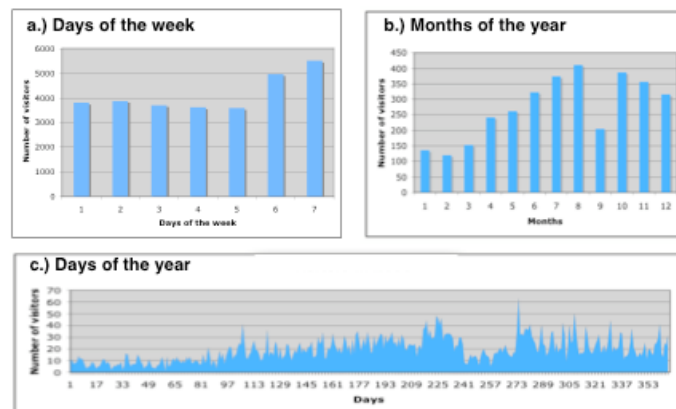


Figure 3. Pulse in average number of visitors for (a) each day of the week, (b) month of the year and

(c) day of the year

Based on the 2006 data, we further went on developing animations of monthly activity, to understand where and when was taking place the tourist activity (Figure 4). To do so, our software developed on top of the Processing⁴ Java library created a layer on the top of the map, showing 250 x 250 m pixels whose colors were related to the number of photos taken. Similarly to Real-time Rome animations based on Erlang values (Reades et al., 2007), we plotted the number of photos with a minimal threshold of 5 pictures to reveal clustered areas. We smoothen the visualization of areas of activity with the use of an interpolation algorithm. In the future, each cell will contain more extensive information to query such as the nationality, the number of people, the time of detection of arrival in the area and the time of departure.

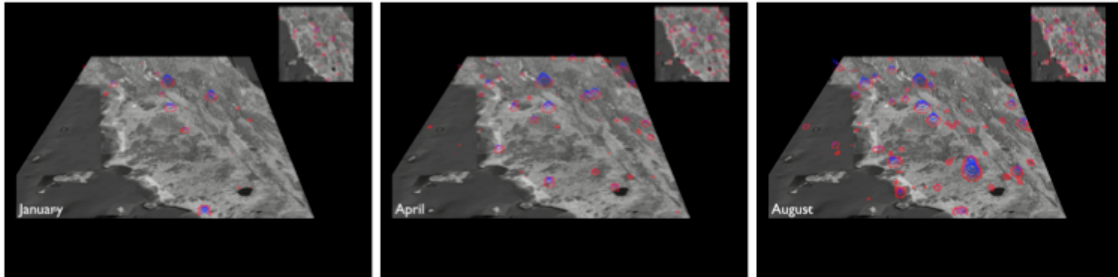


Figure 4. Screenshot of the software animation of the spatio-temporal activity in the Province of Florence.

4.3 Flow into and out of the province

Tourist authorities use classic survey-based hotel and museums frequentation data to get to know where tourists of different nationalities come from and go. However, these data do not capture the “day trippers”, that is the tourists who visit Florence but do not sleep in town. Our data brought a new perspective on the issue of understand where the visitors had been prior to entering the area of study and to where there were heading after leaving the area. We retrieved all the images the 2880 visitors of the area took prior to entering and after leaving the area. In total, we stored the information of more than 1mio of geographically referenced photos taken world-wide. Then, we recovered the inbound and outbound traces from the rules described in Table 2. A specific problem was to give a good estimation of the origin of an inbound movement and the destination of an outbound movement. To solve it, we applied an arbitrary 48 hours threshold between the times of origin and destination of the inbound/outbound traces. The data exceeding this interval were disqualified because they might not have reflected the origin/destination of a day trip in Florence.

Table 2. Description of the algorithm to detect inbound and outbound traces

Type of trace	Algorithm
Inbound	Detect the position and time of the last photo before entering the area, and the position and time of the first photo after entering the area.
Outbound	Detect the position and time of the first photo after leaving the area, and the position and time of the last photo before leaving the area.

We plotted two types of maps to provide information on where tourists are prior to entering the city of Florence and after leaving it: heat maps and flow maps. The first gives a sense of density of inbound origin (Figure 5a) and outbound destination (Figure 5b) of a day trip in

⁴ <http://processing.org/>

Florence.

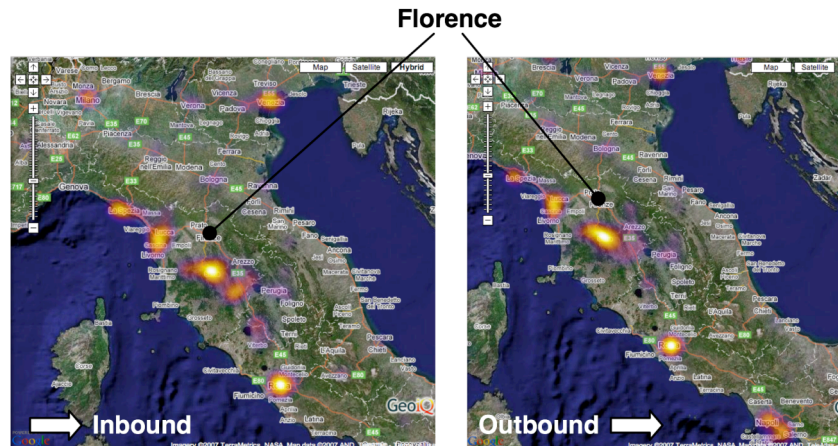


Figure 5. (a) inbound starting point, (b) outbound destination. *Photographic imagery copyright 2007 Cnes/Spot, NASA, DigitalGlobe and TerraMetrics.*

Flow maps offer another perspective on the movement by rendering the streams of tourist. They are generated from KML files of origin and destination traces load in the Google Earth software⁵. For flow leading out of Florence (Figure 6b), each green trace links the position of the last photo taken in Florence with the position of the first photo taken outside of the city within 48 hours. Figure 6a displays the inbound movement to the city of Florence. The results show striking similarities between inbound and outbound flows.

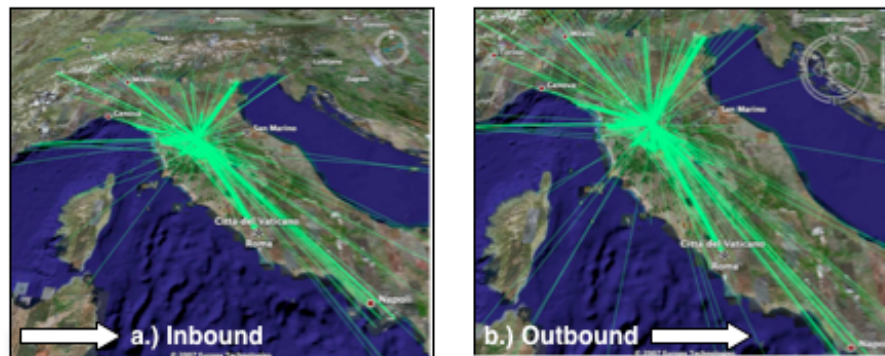


Figure 6. Inbound (a) and outbound (b) flow maps. *Photographic imagery copyright 2007 Cnes/Spot, NASA, DigitalGlobe and TerraMetrics.*

4.4 Patterns of flow

In addition to revealing the origins and destinations of the tourists, the data set contains indications on the flows within the boundaries of the studies area. We could trace the users of Flickr from the digital footprints forming a trace they leave along their path. Practically, a trace consists in a chronologically ordered set of geographically referenced photos taken by one person over the course of one day. Aggregating these personal traces reveal the travel behaviors of specific types of visitors. For instance, Figure 7a reveals that Americans follow a specific graph constituted by the nodes of Florence, Siena, Pisa, Genova and Perugia. On the other hand, Italians (Figure 7b) are more adventurous in their exploration of the area (including reporting on visits of the Island of Elba)

⁵ <http://earth.google.com/>

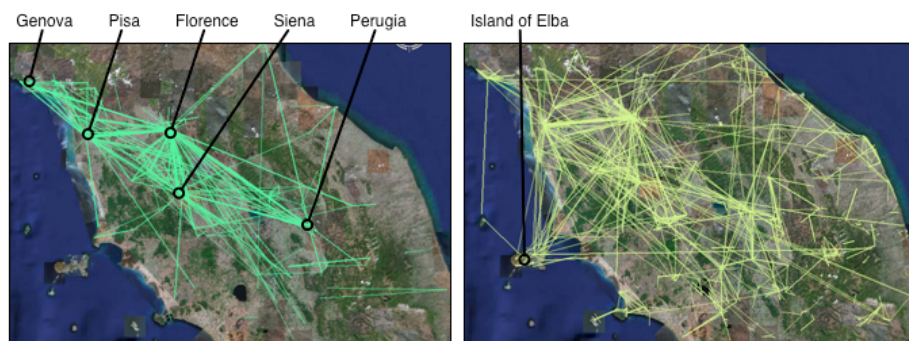


Figure 7: Movements of Americans (left) compared to Italians (right) in northern part of central Italy. Photographic imagery copyright 2007 Cnes/Spot, NASA, DigitalGlobe and TerraMetrics.

In practice, our basic flow maps can easily become too cluttered preventing the emergence of flow patterns. A future solution could be to apply a scale-dependent threshold based on distance and time (e.g. city in meters and few hours, country in kilometers and days). Another solution would be to develop and use a specific flow map tool clustering the main attractions and cities. In any case, what we deem most important in this exploratory phase of research, is to have proved the concept that user-generated content can be used for spatial analysis. From time to time, depending on the given research questions, data sets will have to be interrogated in different ways and visualizations customized.

5 Conclusion and future work

The explosion in the use of captures devices (e.g. mobile phone, digital cameras) and the emergence of content sharing platform led to the emergence publicly available user-generated geospatial data. This document reports on an innovative use of this new type of information to perform urban and mobility analysis. The benefit of our solution lies on the analysis of publicly available individual traces, while most other travel survey and urban sensing infrastructure are limited to privacy-sensitive or aggregated information preventing to trace the behavior in a urban space. Our proof of concept took the context of the use of Flickr to geographically reference photos in the Province of Florence in Italy. Our preliminary results show that the mapping of these data could open a new perspective in urban cartography and lead to a new urban paradigm based on the analysis of publicly available citizen generated geo-referenced data. The explicit act of sharing this information with the ability to control and cloak the data greatly reduces privacy concerns inherent to current travel surveys techniques. In the future, similar types of spatially anchored user-generated content might surface to become relevant for travel, mobility and urban studies. Sources range from implicit data from the usage of radio-frequency networks (GPS, GSM, WiFi, Bluetooth) to more explicit information contained in geostapial web applications (e.g. geo-referencing in Flickr), and the emerging social applications based on the disclosure of presence and location information. In consequence, urban planners and local authorities might consider the aggregate analysis of these different channels. It must be said, though, that - if user-generated data were to be used for statistical purposes - special attention would have to be placed on the issue of self-selection in the data sample. Indeed, those people who proactively upload digital information on open platforms such as Flickr are very likely to represent a tech-savvy subsection of any given population. By collecting further data on the socio-demographics of these users, it will be possible to better understand how much can be generalized as for the behavior of a given population from research findings based on a self-selected sample.

As part of further work, we consider several extension to the promising results presented in this paper. First we aim at correlating our data set with other spatio-temporal data such as GSM

network usage, hotel and attractions surveys. Second, we believe that the recording of such data could be used to inform on the design and deployment of location-based services to enhance the tourist experience. Indeed, so far location-aware application have thus far tended to concentrate on using a mobile carrier's immediate geographic location in isolation but there is a merit in using a position history to tailor results from requests for information further (Mountain and Raper, 2001). For instance, understanding where, when and with which granularity visitors of a city disclose information could improve the relevance of a location-based service aimed at tourists.

Acknowledgements

The results reported in this paper are part of a broader research led in the Wireless City, Florence and Beyond project⁶. We would like thank the authorities of the Province of Florence for their support. Also, we are indebted to many people at the Massachusetts Institute of Technology and the Pompeu Fabra University for providing extremely stimulating research environments and for their generous feedbacks. In particular, thanks to Assaf Biderman, Francisco Calabrese, Bernd Resch and Toni Navarrete for letting us pick their brains. Of course, any shortcomings are our sole responsibility.

References

- Ahern, S., Naaman, M., Nair, R., and Yang, J. H. (2007). World explorer: visualizing aggregate data from unstructured text in geo-referenced collections. In *JCDL '07: Proceedings of the 2007 conference on Digital libraries*, pages 1–10, New York, NY, USA. ACM Press.
- Asakura, Y. and Iryob, T. (2007). Analysis of tourist behaviour based on the tracking data collected using a mobile communication instrument. *Transportation Research Part A: Policy and Practice*, 41(7):684–690.
- Ashbrook, D. (2002). Learning significant locations and predicting user movement with gps. In *ISWC '02: Proceedings of the 6th IEEE International Symposium on Wearable Computers*, page 101, Washington, DC, USA. IEEE Computer Society.
- Eagle, N. and Pentland, A. S. (2006). Reality mining: sensing complex social systems. *Personal Ubiquitous Comput.*, 10(4):255–268.
- Hazas, M., Scott, J., and Krumm, J. (2004). Location-aware computing comes of age. *IEEE Computer*, 37(2):95–97.
- Kracht, M. (2004). Tracking and interviewing individuals with gps and gsm technology on mobile electronic devices. In *Seventh International Conference on Travel Survey Methods*.
- Krumm, J. and Horvitz, E. (2006). Predestination: Inferring destinations from partial trajectories. In *UbiComp*, pages 243–260.
- Mountain, D. and Raper, J. (2001). Modelling human spatio-temporal behaviour: a challenge for location-based services. In *6th Internat. Conference on GeoComputation*. University of Queensland, Brisbane, Australia.
- O'Neill, E., Kostakos, V., Kindberg, T., gen. Schieck, A. F., Penn, A., Fraser, D. S., and Jones, T. (2006). Instrumenting the city: Developing methods for observing and understanding the digital cityscape. In *UbiComp*, pages 315–332.
- Ratti, C., Pulselli, R. M., Williams, S., and Frenchman, D. (2006). Mobile landscapes: Using location data from cell-phones for urban analysis. *Environment and Planning B: Planning and Design*, 33(5):727 – 748.
- Reades, J., Calabrese, F., Sevtsuk, A., and Ratti, C. (2007). Cellular census: Explorations in urban data collection. *IEEE Pervasive Computing*, 6(3):30–38.
- Rattenbury, T., Good, N., and Naaman, M. (2007). Towards automatic extraction of event and place semantics from flickr tags. In *SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information*

⁶ <http://senseable.mit.edu/florence/>

retrieval, pages 103–110, New York, NY, USA. ACM Press.

- Schoenfelder, S., Axhausen, K., Antille, N., and Bierlaire, M. (2002). Exploring the potentials of automatically collected gps data for travel behaviour analysis - a swedish data source. In Möltgen, J. and Wytzisk, A., editors, *GI-Technologien für Verkehr und Logistik*, number 13 in IfGIprints, pages 155–179. Institut für Geoinformatik, Universität Münster.
- Snavely, N., Seitz, S. M., and Szeliski, R. (2006). Photo tourism: exploring photo collections in 3d. *ACM Trans. Graph.*, 25(3):835–846.
- Sohn, T., Varshavsky, A., LaMarca, A., Chen, M. Y., Choudhury, T., Smith, I., Consolvo, S., Hightower, J., Griswold, W. G., and de Lara, E. (2006). Mobility detection using everyday gsm traces. In *UbiComp*, pages 212–224.
- Stopher, P., Bullock, P., and Jiang, Q. (2003). Visualising trips and travel characteristics from gps data. *Road & Transport Research*, 12(2):3–14.
- Torniai, C., Battle, S., and Cayzer, S. (2007). Sharing, discovering and browsing geotagged pictures on the web. Technical report, HP Labs.
- Wermuth, M., Sommer, C., and Kreitz, M. (2003). *Impact of New Technologies in Travel Surveys*, chapter 27, pages 455–482. Pergamon, Oxford.
- Wolf, J. (2004). Applications of new technologies in travel surveys. In *7th International Conference on Travel Survey Methods, Costa Rica*.
- Wolf, J., Guensler, R., and Bachman, W. (2001). Elimination of the travel diary: experiment to derive trip purpose from global positioning system travel data. *Transportation Research Record*, (1768):125–134.