

Tuesday, June 15, 2010

Flickr Photos Yield Tourist Trails

Software uses images from millions of tourists to suggest ways for visitors to spend their time.

By Tom Simonite

Tourists pondering how to spend their time in a large city could one day get help from a tool developed by researchers at Yahoo. It draws on the database of millions of photos uploaded to the site Flickr to generate detailed itineraries of what sites to visit, and in what order.

"If I have two days to fill in Rome, say, the current state of the art is that I have to spend hours reading websites like [Yahoo Travel \(http://travel.yahoo.com/\)](http://travel.yahoo.com/) and looking at online maps to come up with even a simple itinerary to fit the time I have," says [Munmun De Choudhury \(http://www.public.asu.edu/~mdechoud/\)](http://www.public.asu.edu/~mdechoud/), who worked on the project with colleagues from Yahoo's research labs in New York and Haifa, Israel, and is now at Arizona State University, in Tempe. "We thought, why not use the crowd's wisdom?"

The tool works for five cities: Barcelona, London, New York, Paris, and San Francisco. To inform its suggestions, it extracted tourists' movements between attractions in those cities from millions of photos uploaded to Flickr over three years.

Geolocation data and the tags added by users were used to determine and locate which attractions each user had visited, with reference to lists of attractions in the cities sourced from Yahoo Travel and Lonely Planet; the timestamps revealed how much time was spent at each attraction and how long it had taken to travel between them. Snaps from nontourists were excluded by only considering photos covering a span of time in a city shorter than a few weeks. Although the dataset used is not public, it would be possible to reproduce everything the tool does using the [Flickr public application programming interface \(http://www.flickr.com/services/api/\)](http://www.flickr.com/services/api/), says De Choudhury.

De Choudhury and colleagues developed a simple Web interface that can be used to automatically generate an itinerary to fill a specific number of days in a city. The Flickr data is used to fill the time available with visits to the most popular attractions and to specify the amount of time that should be spent at each. The order of visits is chosen to

minimize travel time.

The quality of the results was tested by more than 450 users of Amazon's [Mechanical Turk](https://www.mturk.com/mturk/welcome) (<https://www.mturk.com/mturk/welcome>) site--a platform for paying large numbers of people to perform relatively simple tasks. To ensure that the people doing the judging knew something of the city in question, a screening step asked them to identify a photo of a relatively minor attraction from each city. For example, a photo of [Pont Neuf](http://en.wikipedia.org/wiki/Pont_Neuf) (http://en.wikipedia.org/wiki/Pont_Neuf) was shown for Paris.

One test presented each participant (known as a "turker") with two itineraries: one generated from the Flickr photos, and one sourced from a professional tour operator based in the same city. "Most people actually found our itineraries better than the professional ones," says De Choudhury. "Seventy percent said that ours were significantly or somewhat better."

A second test presented a single automatically generated itinerary alongside a professional one and asked a series of questions about the appropriateness of the suggested attractions, the time to be spent at each, and transit times between them. People scored both itineraries roughly the same.

[Fabien Girardin](http://www.girardin.org/fabien/) (<http://www.girardin.org/fabien/>), cofounder of [Lift Lab](http://www.liftlab.com/) (<http://www.liftlab.com/>), a technology research agency based in Switzerland, has also worked on extracting the movements of people from Flickr photos, using the results to evaluate the popularity of public artworks in New York. "They formalize this approach a little better and go full circle by actually starting to build what can be extracted into information that people can use," he says of the Yahoo researchers' work.

It should be possible to go further still, Girardin adds, by generating more interest-specific itineraries. "They talk about the general tourist, but there are many ways to see a city."

De Choudhury says that adding personalization is a logical next step. "The current itineraries might be good if this is the first time you're going somewhere, but of course some people might want to stay in galleries all day, or concentrate on historical sites," she says. "Those kinds of personalizations are possible."

Ultimately, though, says Girardin, not many people are likely ever to follow such directions to the letter. But that doesn't mean they are not useful. "It's like a weather report," he says. "You combine the information with other things you see around you."

Copyright Technology Review 2010.

Upcoming Events

[EurekaFest \(http://www.eurekafest2010.org\)](http://www.eurekafest2010.org)

MIT Stata Center; Museum of Science, Boston

Wednesday, June 16, 2010 - Saturday, June 19, 2010

<http://www.eurekafest2010.org> (<http://www.eurekafest2010.org>)

[Tech Connect World \(http://www.techconnectworld.com\)](http://www.techconnectworld.com)

Anaheim, CA

Monday, June 21, 2010 - Friday, June 25, 2010

<http://www.techconnectworld.com> (<http://www.techconnectworld.com>)

[2010 IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply \(http://www.ieee-energy.org/\)](http://www.ieee-energy.org/)

Waltham, Massachusetts

Sunday, September 27, 2009 - Tuesday, September 28, 2010

<http://www.ieee-energy.org/> (<http://www.ieee-energy.org/>)

[FutureM \(http://www.futurem.org/Default.aspx\)](http://www.futurem.org/Default.aspx)

Boston, MA

Monday, October 04, 2010 - Friday, October 08, 2010

<http://www.futurem.org/Default.aspx> (<http://www.futurem.org/Default.aspx>)