

# Finding Wormholes with GeoTags

Maarten Clements  
Arjen P. de Vries & Marcel J.T. Reinders

## Introduction

Many visitors of travel websites like [www.expedia.com](http://www.expedia.com) have not chosen a location for their holiday when they enter the website yet. To actively assist indecisive users, these websites show cheap travel deals. Knowledge about similar travel locations can boost the sales by showing relevant locations for specific users or queries.

## Flickr GeoTag data

To study people's travel behaviour we have collected the geotags of 36,264 users from [www.flickr.com](http://www.flickr.com). Together these users have uploaded 52,425,279 photos of which 22,710,496 have been geotagged.

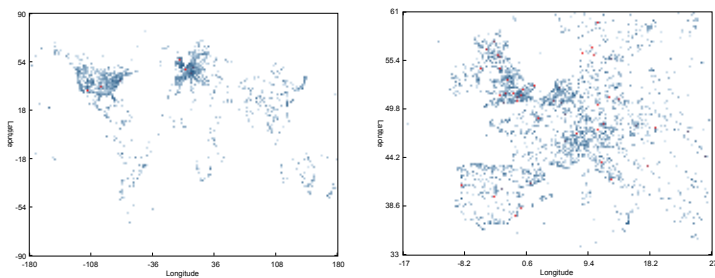


Figure 1. Histogram of the geotags collected from Flickr

## Finding Wormholes

From a given target location  $L$  we want to find the most similar locations around the world. For each user  $u$ , a weight  $W_u$  is computed using the standard normal distribution:

$$W_u = \exp\left(-\frac{\min_i(d(L, G_{u,i}))^2}{\sigma^2}\right), \quad (1)$$

where standard deviation  $\sigma$  is used as a scaling parameter and  $d(L, G_{u,i})$  computes the euclidean distance between one of the user's geotagged images  $i$  and  $L$ .

The wormholes from  $L$  are now derived by creating a 1000x1000 histogram of all users' geotags, using  $W_u$  as weight per user and convolving this distribution with a Gaussian kernel. The difference between the unweighted distribution and the predicted profile indicates whether the relevance probability of a location is above or below the expected value.

## Wormholes at Mount Everest

When Mount Everest is selected as target location, the predicted wormholes should take us to other mountains around the world.

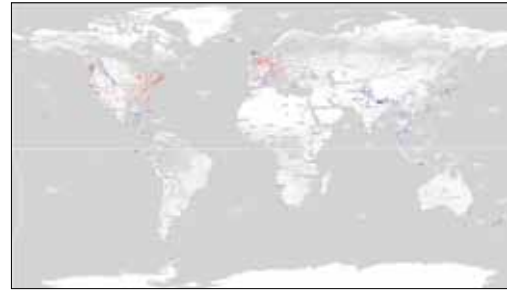


Figure 2. The predicted wormholes from Mount Everest ( $\sigma = 50$  km)  
Positive holes are blue, negative holes are red

A ground-truth dataset for mountaineers is collected by selecting all mountains listed on [http://en.wikipedia.org/wiki/List\\_of\\_peaks\\_by\\_prominence](http://en.wikipedia.org/wiki/List_of_peaks_by_prominence).



Figure 3. The most prominent mountains in the world

All mountains from Figure 3 are used as test locations. Figure 4 ranks the test points by predicted relevance. Clearly most of the test point can be reached through the predicted wormholes.

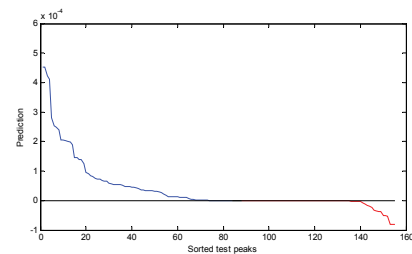


Figure 4. Sorted prediction at test locations



Maarten Clements, M.Sc.

D-MIR Lab, Faculty of EEMCS, Delft University of Technology (TUD)  
Mekelweg 4, 2628 CD, Delft, The Netherlands  
Tel: +31 (0)15 2788612 / Mail: [m.clements@tudelft.nl](mailto:m.clements@tudelft.nl)  
Web: <http://ict.ewi.tudelft.nl/~maarten>

[I,C)<sup>T</sup>

  
**TU Delft**