# A new geographical data source: geo-located Twitter data

Abstract

The increasing

1. Introduction

In particular, Twitter offers

There are several ways to

RESTful APIs

Data

Timzone adjustments

When using Twitter Streaming API to collect geo-located data, the timestamps associated with the data are the local timestamp of the whatever machine the crawler is using.

Geo-located Twitter data is

2. Data Collection

Geo-located Twitter data can be collected in two ways

Geo-located tweets are tweets appended with an additional geo-tag in the form of a pair of geographical coordinates, which represents the location a tweet was sent at.

3. Data processing

3. Case Study

4. Best Practice

5. Conclusion

In this study, the geo-located tweets were downloaded using the Twitter Streaming API[[1]](#footnote-1), where we specified a geographical bounding box as an area-of-interest to retrieve all the geo-located tweets that fall within it. To ensure the complete coverage over Great Britain, we have set the bounding box to British Isles (lower left coordinates in (latitude, longitude): (49.497, -14.854) and upper right coordinates: (61.186, 2.637)), which also includes the whole area of Ireland and a part of France. We have implemented a data crawler and continuously collected 7-month data (1st June – 31st December, 2014) with over 101.8 million tweets and 60 GB in size. During the data collection phase, the data crawler did not encounter any issue regarding whether it exceeds the data quota by the 1% policy mentioned in (Hawelka et al., 2014). It means we have managed to download all the geo-located tweets for the given bounding box. In particular, to showcase the overall spatial coverage of the collected geo-located tweets, a density map of the Twitter user locations in British Isles for July 2014 is shown in Figure 4, where the collected points alone visually reveal the geography of cities and countries, e.g. the clusters with high density of tweets (in color red) correspond to the skeleton of major cities.

1. https://dev.twitter.com/overview/api [↑](#footnote-ref-1)