Mapping of Tweets

P. Zachhuber (e0527176)[1]

[1]**Student** at the Vienna University of Technology, Austria

Abstract

Aim of this Project is, to map short messages published via Twitter. To do so, those messages (tweets) will be saved locally and visualised.

1   How Twitter works

1.1   About Twitter

Twitters Headquarter is based in San Francisco and twitter.com online since January 2000. (1)

Twitter offers a system to publish messages with a length of 140 characters or less. At first it was intended as a short messaging service but then began to grow and started to use other platforms for its service. (2)

1.2   Twitter.com

In the pictures below the daily pageviews are shown. Fig. 1 shows the steady increase of Pageviews over the last year. A comparison with Fig. 2 reveals that the average user visits twitter.com around six times a day. So, maybe the continuing increase as depicted in Fig. 1 might come from unique visitors following links from social networks like e.g. facebook, where around 11% of users surf bevore they come to twitter. (3)
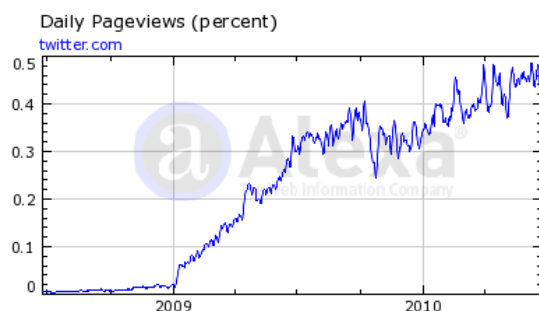

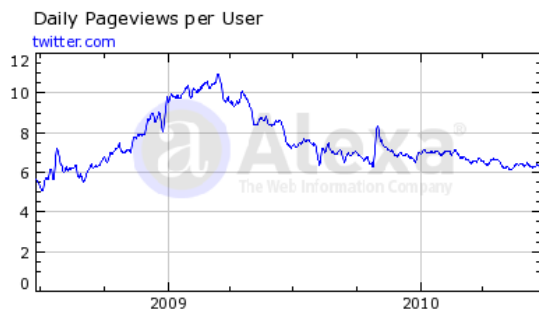Fig. 1: Percent of global pageviews on twitter.com. (3)

Fig. 2: Daily pageviews per user for twitter.com. (3)

1.3   The Twitter API

The Twitter Application Programming Interface (API) is freely available. In 2009 the possibility to add information about the location to tweets started. Unfortunately the possibility to do so is not rolled out in every country by now.

```
"At this time Tweet With Your Location is
only available in the United States. We're
working  on  rolling  it  out  to  other
countries as fast as we can!"

                            'Eddie' (4)
```

Nevertheless, some third-party applications exist for adding positions.

2   This Project

2.1   Aim of the Project

Aim of this Project is, to map short messages published via Twitter. To do so, those messages (tweets) will be saved locally and visualised.

3   System Setup

3.1   The local part

The local part consists of a program that fetches the public timeline[1] and looks for users that post tweets with coordinates. The public timeline is updated every minute and consists of maximal 20 random tweets.

---

[1] http://api.twitter.com/1/statuses/public_timeline.xml

The tweets posted with this identification number are afterwards retrieved from the user-specific timeline[2] and uploaded into the database.

If the IP-address of the requesting device is not explicitly whitelisted those requests are subject to an hourly limit. Because of the uncertain future of this project whitelisting was not applied for.

3.2   Webinterface

The webinterface was realised using HTML with CSS for the layout and PHP for creating dynamic content and database-queries.

3.3   Setting up MySQL

- ### Database

A database named 'twitter' was created within MySQL. This was done using the browser-interface via http://localhost/phpmyadmin/index.php but could also be done using:

```
CREATE DATABASE `twitter` DEFAULT
CHARACTER SET utf8 COLLATE
utf8_general_ci;
```

- ### Tables

Two tables within the new 'twitter'-database were created. The main table consists the tweets with basic fields (shown as XML-structure):

```
<tweets>
    <id>9980714046</id>
    <user>Wisham Faroog</user>
    <userid>14556492</userid>
    <tweet>lorem ipsum</tweet>
    <lat>4.1771299</lat>
    <lon>73.518306</lon>
    <time>2010-03-04 16:40:43</time>
    <geotag>0</geotag>
    <picurl>http ... </picurl>
    <usrurl>twitter.com/...</usrurl>
</tweets>
```

A second table consists names of cities with their corresponding coordinates. This table is here to provide content for a drop-down box. It is also growing as users type placenames into an input field, because this content is geocoded using Googles API and inserted into the locations-table.

```
<locations>
    <City>Vienna</City>
    <lat>48.209206</lat>
    <lon>16.372778</lon>
</locations>
```

4   Used Software

4.1   Python

Python is an open source programming language (5) and comes for example along with ESRI´s ArcGIS. I chose this language because i already used it before and is supported by a lot of well documented external libraries.

4.2   Python Libraries

- ### MySQL-python

This library offers connectivity to MySQL and is therefore an essential part of the written software. http://sourceforge.net/projects/mysql-python/files/

- ### Googlemaps

This library offers easy access to map-services. http://pypi.python.org/pypi/googlemaps/1.0.2

4.3   mySQL

The free 'MySQL' database is used for storing the collected data. At the time of the Project Version '5.1.46-community' was available.

4.4   Gmapper 1.2

Gmapper is an php-class for easier handling of googlemaps within a webpage. It offers easy wrappers fort he googlemaps API.

http://gmapper.ajax-info.de/

5   Acknowledgements

5.1   I would like to thank...

---

[2] http://api.twitter.com/1/statuses/user_timeline/22256565.xml

6   References

1. about. twitter. [Online] [Cited: 28 June 2010.] http://twitter.com/about.

2. McFedries, Paul. twitter - Tips, Tricks, and Tweets. Indianapolis : Wiley Publishing, Inc., 2009. ISBN: 978-0-470-52969-0.

3. Twitter.com Site Info. Alexa - The Web Information Company. [Online] [Cited: 28 June 2010.] http://www.alexa.com/siteinfo/twitter.com.

4. 'eddie'. Twitter Support : About the Tweet With Your Location Feature. Twitter. [Online] [Cited: 28 June   2010.] http://twitter.zendesk.com/forums/26810/entries/78525.

5. python. python. [Online] [Cited: 13 May 2010.] http://www.python.org.