DMSN Coursework 1: Crawling Twitter

The aim of this task is to make use of publically available web APIs to gather social media data. In this instance, you will set up authorisation to Twitter's REST APIs, and query timeline, profile, followers, and friends of a populated Twitter account through it. The response data is JSON-formatted, and can as such easily be used for later analysis.

Twitter App Creation

1. Navigate to: https://apps.twitter.com/app/new and follow the instructions. If you don't have a Twitter account, you will need to register one. Registration should be complete (email and phone number confirmed), otherwise app creation might fail.
2. In Application Management go to the Keys and Access Tokens tab and find the Consumer Key (API Key), Consumer Secret (API Secret). In Token Actions field generate Access Token and Access Token Secret. You will use these in the following steps.
3. For the sake of best practice, set the permissions in the permissions tab to the minimum needed, in this case "Read Only".

Interfacing with the Twitter API

You are not required to interface with Twitter's APIs directly, but can instead use wrapper libraries in the language of your choice. A list of wrapper libraries can be found at https://dev.twitter.com/overview/api/twitter-libraries. It is recommended that you do so using Python 2/3 (using e.g. Tweepy), or Java 5+ (using e.g. Twitter4J). Basic use of these libraries make utilise Application-only authentication (https://dev.twitter.com/oauth/application-only) as opposed to Single-user OAuth (https://dev.twitter.com/oauth/overview/single-user). See docs for differences.

Collecting Data

Using Twitter documentation (https://dev.twitter.com/rest) as a reference, programmatically query four Twitter API endpoints. These can be for your own account, or a different public account with some activity. For this user, query:

- Profile
- Timeline
- Followers
- Followees (aka friends)

Deliverables

A report that includes:

- A screenshot of the application created through Twitters app dashboard (feel free to censor the actual tokens for your own privacy)
- All code used
- A screenshot of JSON data for each endpoint queried above. These do not need to include all the data, but enough to show that the data was collected, and what kind.

Appendix 1: Example Code in Python

1. Install a Tweepy with instructions found here: https://github.com/tweepy/tweepy
2. Copy your tokens from the app dashboard, and use the following code to query your profile and timeline.

```
from __future__ import absolute_import, print_function

import tweepy

consumer_key=''
consumer_secret=''
access_token=''
access_token_secret=''

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.secure = True
auth.set_access_token(access_token, access_token_secret)

api = tweepy.API(auth)

print(api.me())
print(api.home_timeline())
```

Appendix 2: Example Code in Java

1. Download Twitter4J from: http://twitter4j.org/en/index.html (API at: http://twitter4j.org/javadoc/index.html)
2. Copy `lib/twitter4j-core-4.0.4.jar` into the project's folder.
3. Compile the example: `javac -cp twitter4j-core-4.0.4.jar TwitterTest.java`
4. Run the example: `java -cp .:twitter4j-core-4.0.4.jar TwitterTest`

```java
import twitter4j.conf.ConfigurationBuilder;
import twitter4j.*;
import java.util.List;

public class TwitterTest {
    public static void main(String[] args) {
        try {
            // Authorise the library
            ConfigurationBuilder cb = new ConfigurationBuilder();
            cb.setOAuthConsumerKey("Consumer Key (API Key)");
            cb.setOAuthConsumerSecret("Consumer Secret (API Secret)");
            cb.setOAuthAccessToken("Access Token");
            cb.setOAuthAccessTokenSecret("Access Token Secret");

            Twitter twitter = new TwitterFactory(cb.build()).getInstance();
            User user = twitter.verifyCredentials(); // Get main user

            // Print user profile
            System.out.println("@" + user.getScreenName());
            System.out.println(user.getId());
            System.out.println(user.getProfileImageURL());
            System.out.println(user.getFriendsCount() + " friends.");
            System.out.println("-------");

            // Print Home Timeline
            List<Status> statuses = twitter.getHomeTimeline();
            System.out.println("Showing @" + user.getScreenName() + "'s home timeline.");
            for (Status status : statuses) {
                System.out.println("@" + status.getUser().getScreenName() + " - " + status.getText());
            }
        } catch (TwitterException te) {
            te.printStackTrace();
            System.out.println("Failed to get timeline: " + te.getMessage());
            System.exit(-1);
        }
    }
}
```