

Emotwitter – A Tool for Emotional Analysis of Twitter Stream

Yuki Kobayashi

s1200224

Supervised by Prof. Maxim Mozgovoy

Abstract

Social media is rapidly developing now, and its influence is significant for many people. It is used for many purposes, such as communicating with friends and strangers, obtaining information and news, advertising new products and so on. In the University of Aizu we developed a tool called Emotwitter, which can analyze user tweets.

Emotwitter can download tweets of particular users, analyze word frequency, build tweet word cloud and identify emotions contained in tweets. Since the system still needs improvements, this study is dedicated to resolving existing problems and to adding new functions. Improvements include the support of the latest Twitter API 1.1 and OAuth authentication using access token of user account. New functionality includes a keyword-based tweet search function. These improvements enable Emotwitter to provide better analysis options than before.

1 Introduction

On the Internet, different forms of communication are developing. We use various Internet communication tools to get into contact with people. Among them, social media has a significant place. Social media is a scalable media that provides user-generated contents and helps user to connect with each other.

Twitter is one of the most famous micro blogs in the world. According to official Twitter, Inc. information, it has more than 30 million active users per month all over the world. The number of users who visit a website through a link posted on a tweet is more than billion per month [1]. Twitter is also becoming an important tool for marketing, and especially popular among young people. This trend is expected to continue in the future, so we need tools to analyze tweets. Information volume of Twitter is huge, and its content is diverse: individual emotional messages, news, advertisements of new products and so on. Therefore, there is a demand for software instruments that are able to categorize and analyze tweets, helping to use Twitter resources more effectively.

Emotwitter is a tool for emotional analysis of Twitter messages, created at the University of Aizu [2]. It supports several types of text analysis, but in its recent version also suffered from a number of problems. First, it did not support the latest Twitter API 1.1. It caused occasional faults when retrieving user tweets. Next, Emotwitter did not use OAuth authentication with access token of user account. Without OAuth support it is impossible to redistribute the software to an unspecified group of users (because they will use the same limited access token). So I mainly improved these points in my work.

2 Emotwitter in Detail

Emotwitter is a tool written in C# with the support of several third-party libraries to analyze user emotions. The Fig.1 shows the control panel of Emotwitter. When a user inputs a Twitter account name into the text box "Username" and clicks the "Twitter" button, the system displays the retrieved tweets of the given user. In the example it is visualizing the tweet stream of @darthvader¹. Emotwitter can get no more than 3200 tweets using Twitter API (it is a limitation of Twitter API). If a user inputs certain keywords into the text box "Filter by keyword", the system will extract tweets from the list of downloaded tweets that contain the given keyword. Emotwitter can also build the following types of visualization:

- Emotional distribution by calculating emotional score using word-emotion association lexicon.
- Rader chart of emotional score (see the Fig.2).
- Word cloud of frequent words in tweets (see the Fig.2).

¹Darth Vader is a fictional character in Star Wars.

3 Emotwitter Improvements

3.1 Correspond Twitter API 1.1

Emotwitter used Tweetsharp — one of C# libraries that support communication with Twitter. However, Tweetsharp support was finished in 2014, and its final version is not compatible with the latest twitter API 1.1. As a result, Emotwitter operations often failed with no clear reasons.

So I decided to substitute it with the modern Twitter library Tweetinvi [3]. Tweetinvi is a C# library designed to simplify the development of applications for Twitter in C#. When we migrated to this new Twitter library, Emotwitter started working normally. The Fig.3 is a part of the source code used to download tweets. Migration to Tweetinvi also required numerous modifications in this module. Normally, "user timeline" API can download only max 200 tweets per call, so it is necessary to perform

numerous requests. Emotwitter implements using "Max_id" and repeats Fig.3 instruction.

3.2 OAuth Authentication

OAuth is an open protocol that provides algorithms for secure authorization. So it leads saving user password and it can say secure authentication. [4]

OAuth authentication does not use user password, instead it relies on the access token sent by the application. So it leads saving user password and it can say secure authentication.

Each Twitter API-enabled application needs a Twitter account to rely on consumer key, consumer secret sequence, access token and access token secret sequence of a Twitter account. Consumer key and consumer secret sequence must belong to the developer, but the access token and the access token secret sequence should belong to the application user (since Twitter limits the number of requests per user).

```
var oldestTweet = tweets.Min(x => x.Id) - 1;
var userTimelineParameter = new UserTimelineParameters();

userTimelineParameter.MaxId = oldestTweet;
userTimelineParameter.MaximumNumberOfTweetsToRetrieve = 200;

receivedTweets = Timeline.GetUserTimeline(userName, userTimelineParameter).ToArray();
tweets.AddRange(receivedTweets);
```

Figure 3: Retrieving tweets

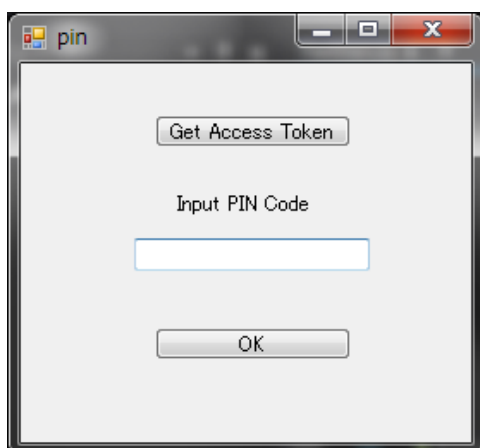


Figure 4: When user starts Emotwitter, first this window opens



Figure 5: The upper part is the authentication page. The lower part is a PIN code.

system for identifying enduring sentiments in tweets,” in Computational Linguistics and Intelligent Text Processing, pp.78–91, Springer, 2015.

[3] Tweetinvi, “<https://tweetinvi.codeplex.com>.”

[4] 太田侑介, “Twitter におけるリツイート経路の可視化,” 電気通信大学情報通信システム学講座, p.13, 2011

```
var searchParameter = new TweetSearchParameters(userName)
{
    Lang = Language.English,
    SearchType = SearchResultType.Mixed,
    MaximumNumberOfResults = 100,
};
var tweets2 = Search.SearchTweets(searchParameter).ToArray();
var user = User.GetUserFromScreenName(userName);

tweets.AddRange(tweets2);
int requestNum = 10;
for (int y = 0; y <= requestNum; y++)
{
    var oldestTweet = tweets.Min(x => x.Id) - 1;
    searchParameter.MaxId = oldestTweet;

    tweets2 = Search.SearchTweets(searchParameter).ToArray();
    user = User.GetUserFromScreenName(userName);
    tweets.AddRange(tweets2);
}
```

Figure 9: Keyword-based tweet search.