CSCI8380 (Fall 2013): Paper Review Form

Reviewer Name: Ting Xiao

Paper Name: Twitris 2.0: Semantically Empowered System for Understanding Perceptions From Social Data

Section I. Overview

A. Reader Interest

1. Which category describes this manuscript?
   _X_ Practice/Application/Case Study/Experience Report
   ___Research/Technology
   ___Survey/Tutorial/How-To

B. Content

1. Please explain how this manuscript advances this field of research and/or contributes something new to the literature.

   Facing the data deluge of Twitter, authors show an application of Twitris, whose semantics analysis can extract spatially-sensitive facets.
   They mainly introduce a semantic social web approach to detect events from the massive data in the social web. It is composed of 6 steps. Specially, they divide the semantics analysis into two parts, one is deep semantics, and the other one is shallow semantics, which build a good foundation for the whole mining process.

C. Presentation

1. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on?
   _X_ Yes
   ___Could be improved
   ___No

2. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic?
   _X_ Satisfactory
   ___Could be improved
   ___Poor

3. Please rate and comment on the readability of this manuscript.
   _X_ Easy to read
   ___Readable - but requires some effort to understand
   ___Difficult to read and understand
   ___Unreadable

Section II. Evaluation
Please rate the manuscript. Explain your choice.

_Award Quality
__Excellent
_X_Good
__Fair
__Poor

The structure of the paper is clear. However the real social web is still a dynamic environment, the “real time” should be changed into one day before.

**Section III. Detailed Comments** (provide your thoughts/criticism about the ideas in the paper; not only summarize the paper but have a critical look here)

We can see Twitris 2.0 just keeps on tracking the tweets of each day. I have checked the mining result of the “Japan Tsunami”, it just feedback a few tweets about this events. As I known, lots of Japanese like using Twitter to post micro-blog, this big disaster should have much higher traffic analysis as a result of Twitris, but there is no detail result of this events on the web site. Maybe I used it not correctly or this application only extracts the English-Tweets. If so, the application has one big limitation.

Additional Comments:

1. Provide one aspect that you liked the most in this paper.

   The application is close to our life and very useful. It can be used to extract every events published by users and monitor the important events around us, for example, swine influenza, firestorm and so on.

2. Provide one aspect that you disliked the most in this paper.

   I don't like the paper only text, it should provide some good figures to help to illustrate the problem. As we seen, there is one graph used to show the functional overview, but without reading context, you cannot capture how this application works. They even don't need this figure.

**Section IV. Discussion Points** (provide at least 3 discussion topics/questions related to ideas/techniques described in the paper; these will be used for discussions in the class)

1. There must be a lot of statistic methods used when doing the data-mining. In this paper, I paid attention to the two statistic computations: 1 TFIDF, 2 N-gram summaries. Can it be used to analysis other language-tweets? Do you know some other useful statistic methods used in semantic web analysis?
2. How to do the syntactic analysis of URL? This method is used to analysis the internal context.
3. What is the lexicon-based classification algorithm?