

A Linguistic and Statistical Analysis of Twitter Trends

Thomas Ring

Design Goals

Utilizing current technology, we can advance market research. Twitter offers thousands of voices and opinions on an open API for free, which is a much better cost than existing market researching tools. If we can utilize other technologies to analyze the linguistics of twitter we can create ratings for services, TV shows, and products on demand, and in much faster speed.

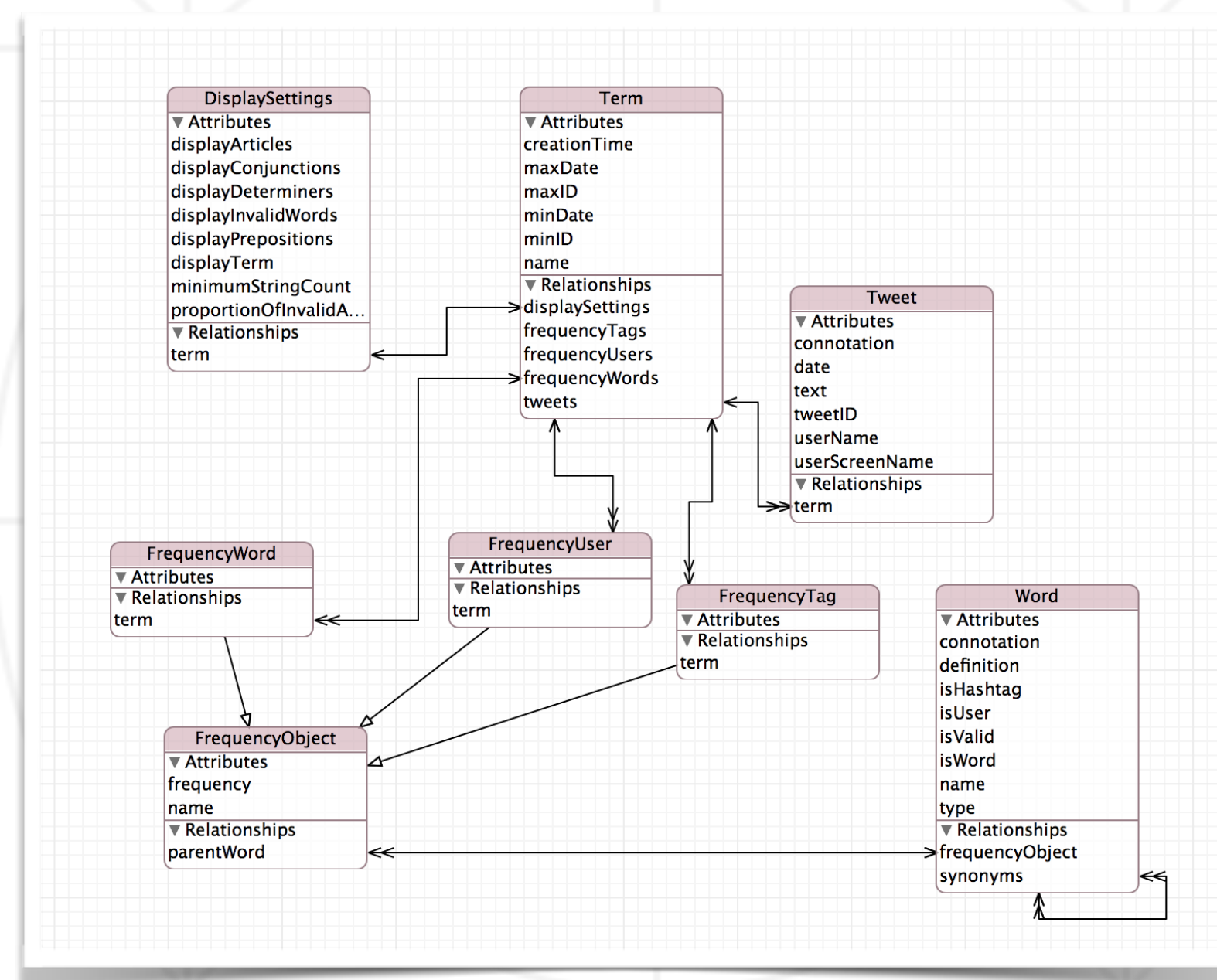
Technologies

Twitter's API provides access to a stream of 58 million tweets per day, limited to ten days, for any application that authenticates with their developer program.

Apple has built Twitter integration into their mobile OS, removing the troubles of authentication into a couple lines of code.

Apple's Core Data technologies also serves as a great medium between an SQLite database and the code.

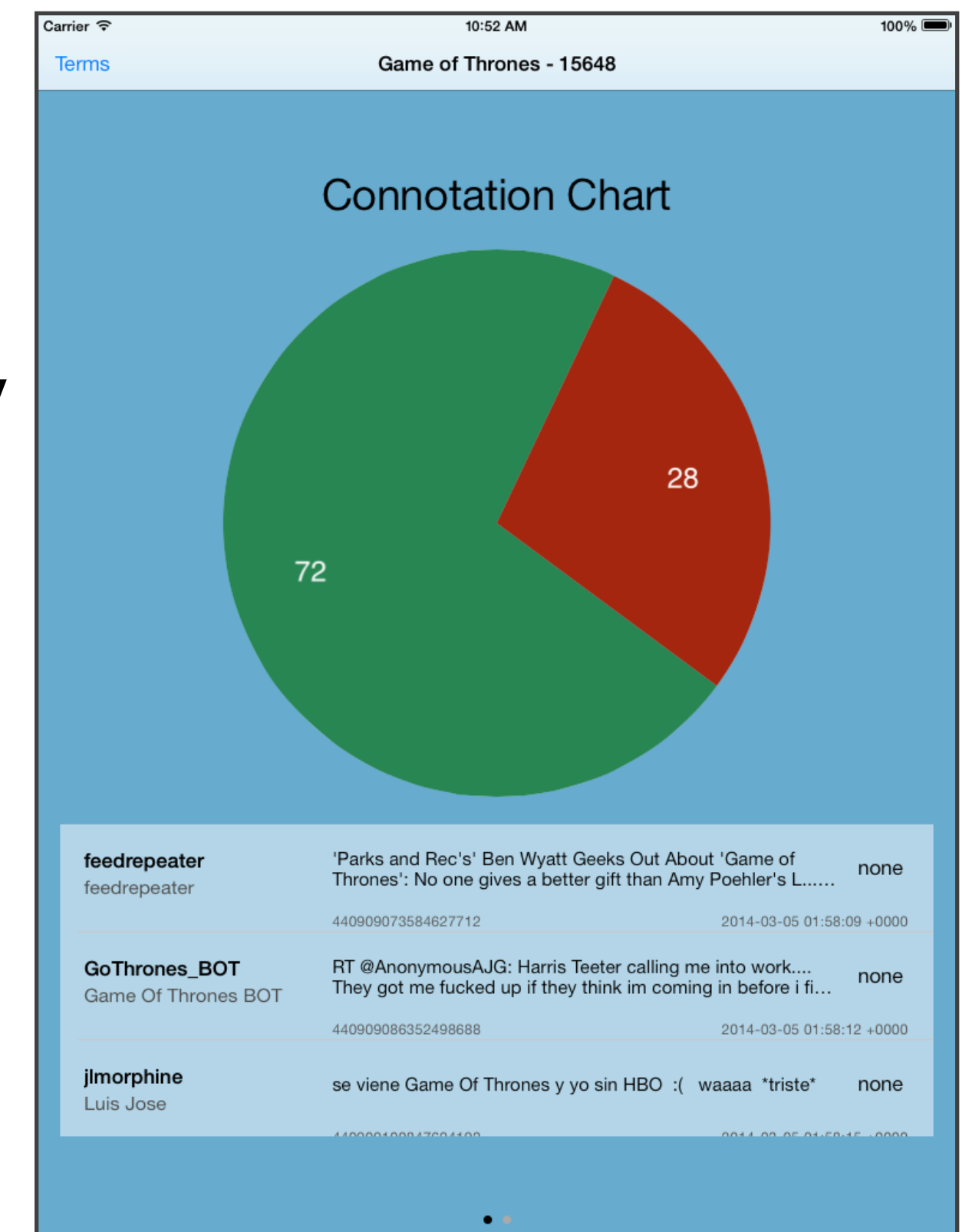
Lexicontext, an offline dictionary for iOS provides information about words like parts of speech, definition and even synonyms to that word.



Current Status

Our Application, code-named TweeterMeter, currently accesses Twitter's API and analyzes tweets based on adjectives used in the sentence for a connotation as "good" or "bad". Data on the frequency of words used is also stored and displayed.

The displays of the frequency of data and proportion of connotation allows for user interaction; when touched, a filtered list of tweets is displayed corresponding to the element touched.



Future Directions

- Improve the connotation algorithm through better dictionary and thesaurus support
- Improve efficiency of tweet sorting at runtime
- Display a timeline graphic for the amount of data is currently stored
- Access Twitter's TweetStream API for realtime results
- iPhone support and Mac OS X support
- Create more aesthetically pleasing graphics