An Image Based Reader for Usenet Headers

Michael Hines
mhines3@illinois.edu
January 29, 2009

Motivation

Anyone who has had to sift through headers in a newsgroup knows that after a while, reading text just gets boring. Text just doesn’t pop out at you the way images do. Unless the subject is “LOSE 15 POUNDS IN JUST 3 HOURS!!!!1” no particular post is likely to immediately grab your attention. It would be nice to browse threads by image, but rather than leave the image up to the person posting, why don’t we find an appropriate image for them?

Method

The visualization will download all the headers from a Usenet group or group on some other similar medium. Then the distinct thread subjects will be sorted in two dimensions, one axis being most recent post date, the other being average message size. Then, a random image will be pulled from Google images based on a search result for a query chosen by one of several methods:

- most frequently used word in the thread (excluding a, the, or, etc.)
- a random noun or verb chosen from the subject
- a combination of words parsed from the first post based on image availability

The images will then be re-sized based on the thread depth. The result will be similar to a tag cloud, with images instead of key words.

Observations

It’s extremely likely that not all of the pictures will match well with the topic of the post. Some will be dead on, and some will be hopelessly off-topic. This is not all bad. If everyone chose to read exactly what they were looking for, how would they be exposed to new ideas? This method allows peoples’ interest to be piqued without necessarily telling them exactly what they are going to read about. After all, chances are that they are interested in the general topic of the group already! Sometimes a little serendipity is exactly what we need to discover something extraordinary.