



TECHNOLOGY

TwitterStand: A Dynamic Aggregator for New Media

OVERVIEW

Researchers at the University of Maryland have built a spontaneous news medium for news generation and collection with the idea of bypassing the traditional setup of reporters and newspapers. The system, called TwitterStand, uses messages (i.e., Tweets) posted in Twitter to automatically obtain breaking news. The key novelty behind TwitterStand is one of mobilizing the millions of users in Twitter to be its eyes and ears in the world, bearing in mind that geographically proximate users often Tweet about the same breaking news. In other words, TwitterStand relies on Twitter users to be either providers of original news content (e.g., the 2008 Southern California earthquake, the 2009 Iranian election, and Michael Jackson's death), or expressers of opinions on current news topics (i.e., mini blogs), both of which enable TwitterStand to automatically identify current news topics and cluster the corresponding Tweets into appropriate news stories. The result is analogous to a distributed news wire service. The differences are that TwitterStand has no reporters being assigned to cover stories. Moreover, the identities of the contributors/reporters are not known in advance and there may be many of them. The Tweets are not sent according to a schedule. They occur as news is happening and are mostly noisy while usually arriving at a high throughput rate.

The challenge lies in identifying, extracting and then differentiating news Tweets from mountains of non-news Tweets, while being able to deal with massive amounts of input data. TwitterStand sorts through Twitter messages in real time, separates newsworthy Tweets from non-newsworthy tweets, and displays them according to their geographic location. TwitterStand makes use of a unique set of news processing algorithms developed at the University of Maryland that robustly cluster Tweets as they occur and readjust the order in which they are presented according to the importance of the news topic. The Tweets that comprise the news topics are presented to users with a map interface thereby enabling them to retrieve Tweets by location. This technology results in a well-organized, attractively-interfaced alternative to news aggregators provided by Google, Yahoo!, Bing, and others. It can be modified or narrowed to extract tweets pertaining to any number of categories including people, diseases, science and technology, health, entertainment, sports, and business.

Applications

- * Spontaneous news medium that completely bypasses conventional news gathering
- * Building block of a sentiment analyzer
- * Universal clipping service (e.g., who is who and where in the news)
- * News application can be modified to fit any number of broad concepts: entertainment, business, etc.

Advantages

- * Map interface for geographically organizing news-related Tweets
- * A faster and more diversely supplied news source than specific online publications
- * Systematically extracts valuable information (news) from noisy Tweets

CONTACT INFO

UM Ventures
0134 Lee Building
7809 Regents Drive
College Park, MD 20742
Email: umdtechtransfer@umd.edu
Phone: (301) 405-3947 | Fax: (301) 314-9502

Additional Information

INSTITUTION

University of Maryland, College Park

PATENT STATUS

Patent(s) pending

LICENSE STATUS

Contact OTC for licensing information

CATEGORIES

- Information Technology

EXTERNAL RESOURCES

IS-2009-116