The OASYS 2.0 Opinion Analysis System: A Demo

Carmine Cesarano,
Antonio Picariello
Dipartimento di Informatica,
Univ. di Napoli Federico II,
Napoli, Italy
cacesara,picus@unina.it

Diego Reforgiato,
VS Subrahmanian
Dept. of Computer Science
and Institute for Advanced
Computer Studies, University
of Maryland
College Park, MD 20742
diegoref,vs@umiacs.umd.edu

1. Background

OASYS 2.0 ("Opinion Analysis System") is a system developed primarly at the University of Maryland in collaboration with the University of Naples, Italy. OASYS 2.0 contains significantly extended and improved versions of algorithms in [1] that allows users to identify the *intensity of opinion* on a -1 to +1 scale on any topic of interest to the user. OASYS allows the user to view how that intensity changed over time, by country, and by news source.

The OASYS 2.0 system currently (as of Dec. 8, 2006) has the following capabilities:

1. Identifies the intensity of opinion expressed by news sources in 8 languages — these languages include English, French, Spanish, Italian, Chinese, Russian, Korean and Arabic.
2. Has analyzed over 4,430,000 news articles (OASYS 1.5 had over 3.1M articles, while OASYS 2.0 has an additional 1.36 million articles at this time).
3. Has analyzed news articles from 109 news sources.
4. Analyzed over 17,700 articles per day.
5. Has analyzed news articles from 19 countries.

OASYS won Computerworld Magazine's 2006 Horizon Awards competition for most innovative pre-commercial software of 2006.

2. What OASYS does: Simple mode

In this section, we briefly describe how a novice user can use OASYS.

OASYS has a crawler that works continuously and crawls the 109 news sites listed above. It continuously analyzes new articles released at those sites and identifies the topics in those articles and assigns a score to the intensity of opinion on the detected topics in that article. This analyzed information is stored in a highly optimized database OpinionDB that we have designed on top of the Postgres system. OASYS can be accessed by any user anywhere in the world using a Java enabled browser.

Figure 1 shows a simple screenshot of how a user uses the simplest features of OASYS (we will highlight a couple of the more advanced features later).

1. The search tab is where the user can type in the search term he is interested in (the search term shown in the search is about Pakistan President Pervez Musharraf).

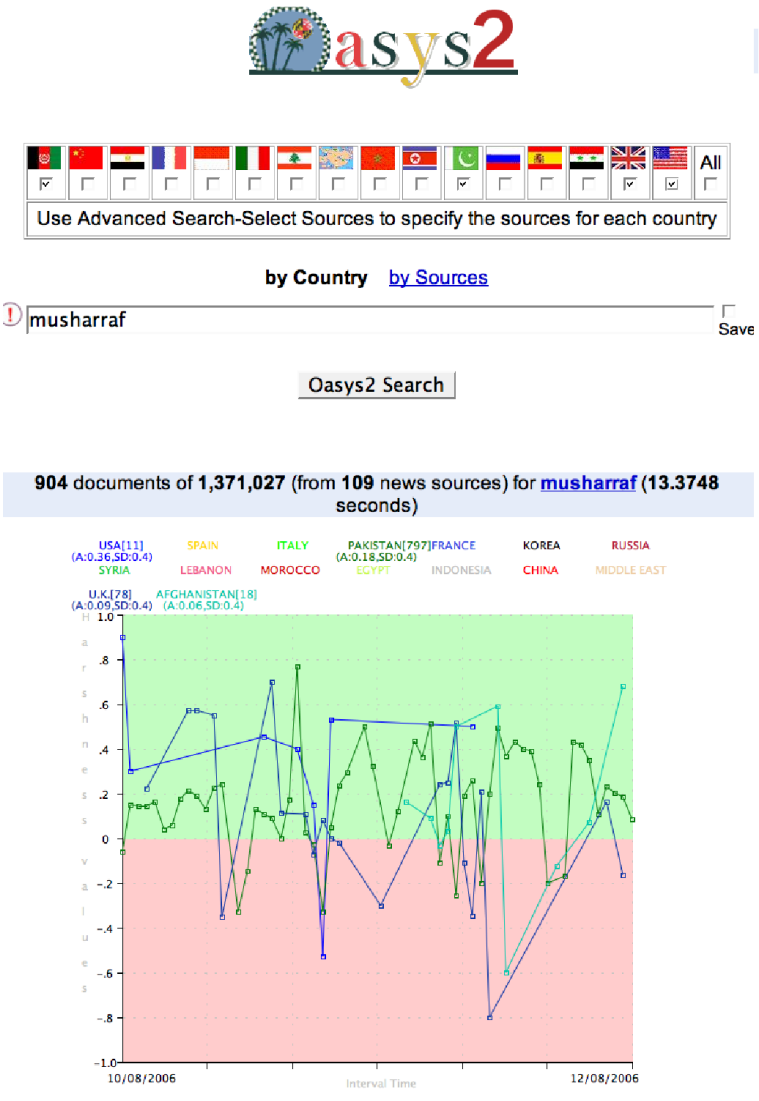


Fig. 1: *Oasys2*

2. The user next selects the country or countries of interest. In the example, the user has checked the US, Pakistan, Afghanistan and UK.

3. Once the user presses the OASYS2 Search button, the system queries the OpinionDB and produces a graph. The graph has time on its x-axis (the default time window considered is the past 2 months). Intensity of opinion is plotted on the y-axis.

For example, in Fig. 1, the reader can see that there are curves plotted in 4 different colors. Each color is associated with a country. The curve shows how intensity of opinion on Musharraf has varied over the 2-month window. In this example, intensity about him has fluctuated a fair amount.

However, in addition to the curves, some statistical information is also provided just above the graph. The string "USA[11], A:0.36, SD:0.4" indicates that there were 11 articles about Musharraf in the USA that we tracked during this time – the average score was 0.36 (which is very high). In contrast, "UK[79],A:0.09,SD:0.4" shows that there was far more reporting about him in the UK (79 articles) and that his rating was much, much lower in the UK than in the USA. Ratings in Afghanistan of President Musharraf were even lower. It should be emphasized that our system only analyzes reported news – so in no way can this be considered to indicate what the common man in these countries think of Mr. Musharraf.

Note that the curves shown in Figure 1 have data points on them. BY clicking on these data points, the user can see the article in its original language, as well as an English translation of it.

3. OASYS use case: Advanced search features

OASYS also has an Advanced Search capability reflected by the button of the same name at the top right corner of Figure 1. When the user selects this option, he has two sets of choices:

1. Select sources allows the user to explicitly select news sources that he wants to track. Pressing this button brings up the list of news sources OASYS can currently track - the user can check those he wants to analyze. This is shown in Figure 2.

2. Select interval time allows the user to override the 2-month default time window and choose the time interval he is interested in.

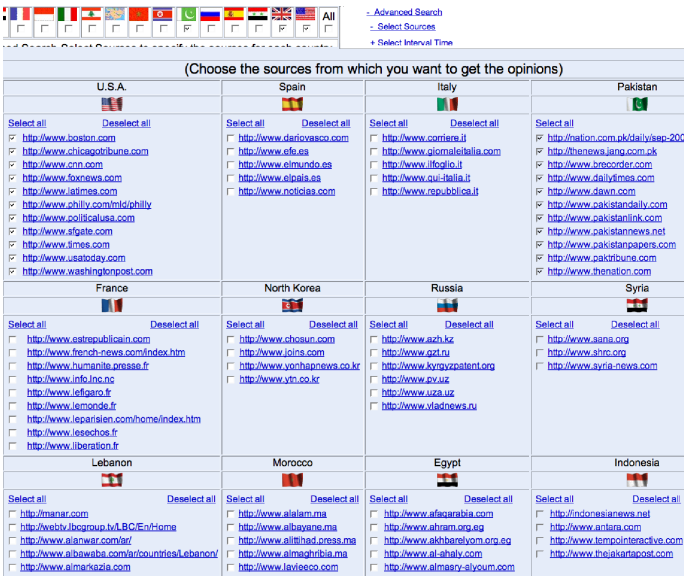
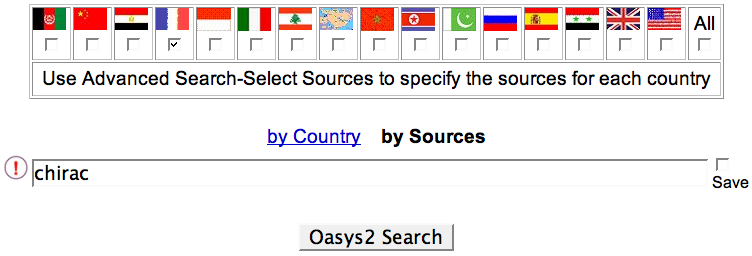


Fig. 2: *Oasys2 Advanced search features*



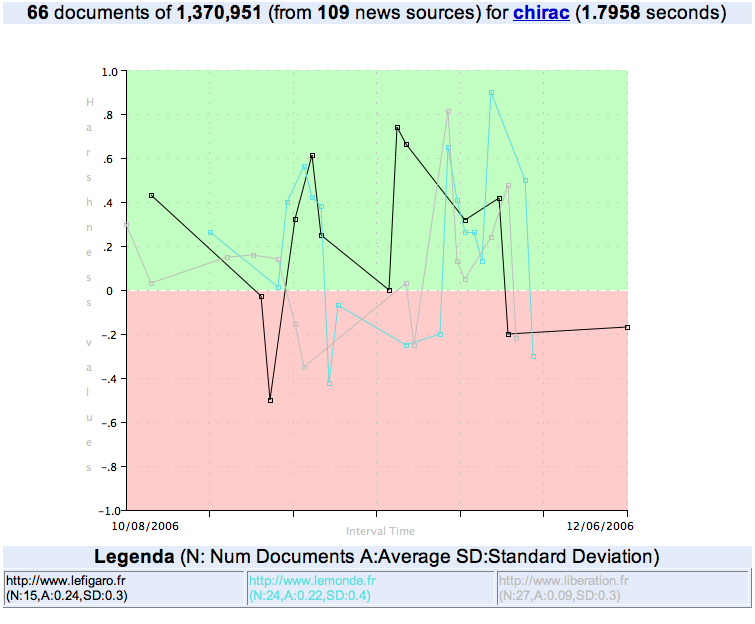


Fig. 3: *Opinions in France about President Chirac*

4. Slicing opinions by news sources

OASYS also allows users to slice opinion information by news sources, rather than by countries. To achieve this, the user presses the By sources button above the search window. Figure 3 shows a sample screenshot of OASYS showing how the French newspapers Le Monde, Le Figaro, and Liberation varied in their opinion of President Chirac over the past two months. Though they all had comparable numbers of articles about Chirac during this time frame, OASYS2 reports that Liberation provided less enthusiastic ratings of Mr. Chirac as compared to the other two sources. This is evident from the averages shown at the bottom of the graph.

Acknowledgment. Work funded in part by AFOSR contract FA95500610405.

References

[1] C. Cesarano, A. Picariello, D. Reforgiato, A. Sagoff and V.S. Subrahmanian. (2006) The OASYS Opinion Analysis System, Proc. 2006 AAAI Spring Symposium on the Analysis of Weblogs, Stanford, CA, March 2006.