Human Sensor Networks: Use of Social Media and Self Organizing Maps for Automated Detection of Oil Spill Plumes in Satellite Observations

by Oleg Aulov, UMBC, Milton Halem, UMBC, Nabin Malakar, UTD, David Lary, UTD

oleg2@umbc.edu, halem@umbc.edu, nabinkm@gmail.com, djl101000@utdallas.edu

UMBC AN HONORS UNIVERSITY IN MARYLAND

The Federation of Earth Science Information Partners — ESIP — Fostering Connections to Make Data Matter

Timely responses to mitigate natural and manmade disasters--particularly deep well oil spills like the Deep Water Horizon oil spill in the Gulf of Mexico in April 2010--can save lives, prevent property damage and help minimize environmental impact. During oil spill disasters, trained satellite analysts at NOAA/NESDIS process satellite observations and manually integrate data from numerous sources to produce a polygonal map that identifies the locations of possible detected oil on the surface of the ocean. These polygon maps are assimilated into an operational Lagrangian trajectory model driven by wind and ocean current data to forecast the movement of the oil. We demonstrate an automated algorithm to detect and map surface oil distributions from satellite observations. We employ a Self Organizing Map (SOM) machine learning algorithm. A SOM algorithm is a type of an unsupervised neural network that produces a low-dimensional representation of a higher dimensional input space while preserving its topological properties. This low order representation is called a map. Once the map is created we use social media data from human sensor networks together with other ground observations to determine which cluster represents the oil plume.

We are viewing social media data gathered from Flicker as geophysical, geolocated, time-stamped data. We view Flickr users as remote sensors, and their posted photos as sensor

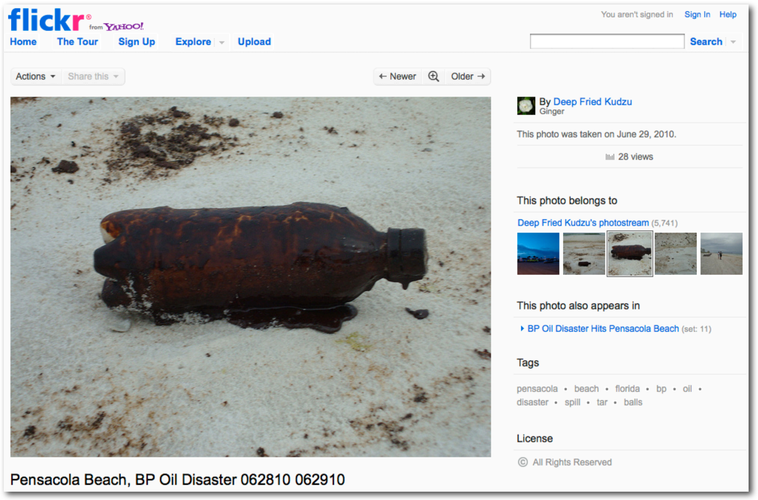

Figure 1

observations. The entire Flickr community is viewed as a human sensor network. Flickr is a website that facilitates its users to share and embed personal photographs and videos and form an online community. Figure 1 shows an example of a Flicker photo of a plastic bevarage bottle covered in oil.

```
API Query:
http://api.flickr.com/services/rest/?method=flickr.photos.
search&api_key=70920bca63b7452f4ff6a7bdbb7f3f75&
tags=tar+balls&
min_taken_date=2010-04-20+00%3A00%3A00&
max_taken_date=2010-10-20+00%3A00%3A00&
bbox=-95.668945%2C+28.07198%2C+-85.825195%2C+31.203405&
has_geo=1&extras=geo%2C+path_alias%2+date_taken&
auth_token=72157626247331502-22c36f6dd37efa88&api_sig=8907818
1e2169738d46489f44f070843

API Response:
<photo id="5016704044" owner="37281343@N03"
secret="9db693ebb5" server="4144" farm="5" title="DSC_1225"
ispublic="1" isfriend="0" isfamily="0" latitude="30.371133"
longitude="-86.918726" accuracy="16"
place_id="uwvhGpebBZlHgiRz" woeid="2457354"geo_is_family=
"0" geo_is_friend="0" geo_is_contact="0" geo_is_public="1"
pathalias="mmmeeks" datetaken="2010-08-04 17:48:24"
datetakengranularity="0" />
```

Figure 2

Flickr API allows programmatic automated gathering of images as well as the metadata associated with those images such as date taken, title, description etc. The photos that were taken with a camera equipped with a GPS, such as an iPhone or an Android smartphone, will include Latitude, Longitude and accuracy as part of it's metadata. Figure 2 shows an example of an API call with an API response containing geolocation and timestamp metadata. Figure 3 shows the map of the affected area with social media data verifying oil observation represented by red dots.

Figure 3

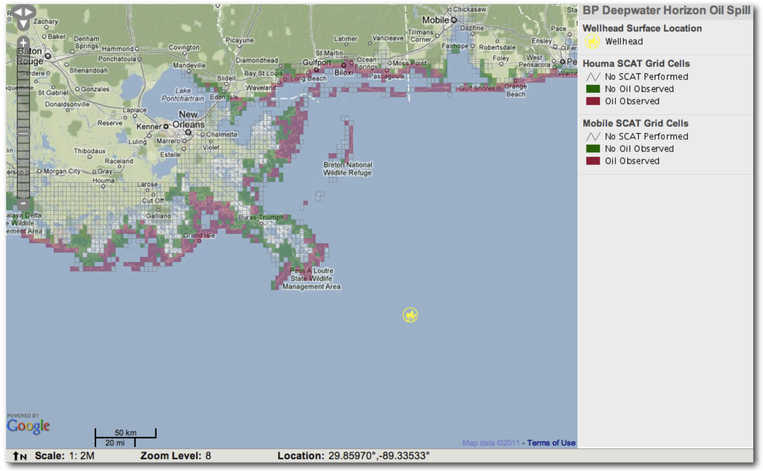

Figure 4

Shoreline Cleanup Assessment Technique (SCAT) program had multiple teams operating across the coastal zones surveying shorelines, making assessments of the oiling conditions and producing a consistent and standardized data collection. Figure 4 shows ERMA web tool displaying SCAT data on the map.

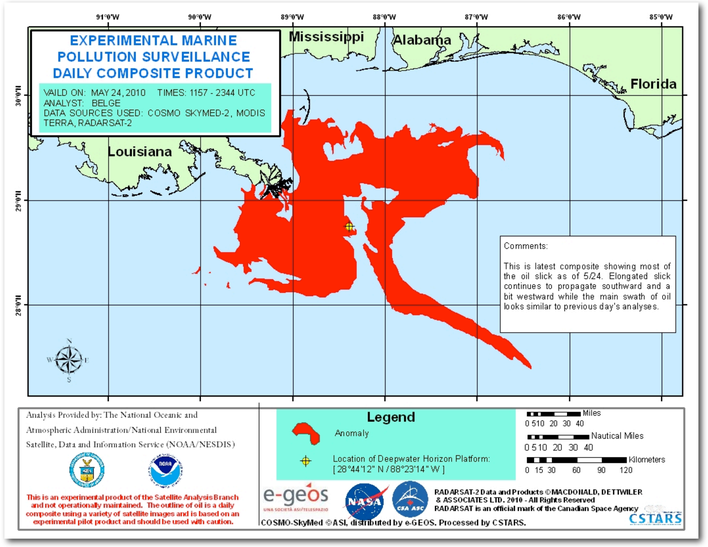

Figure 5

The image on figure 5 shows a NESDIS composite analysis of COSMO SKYMED-2, MODIS TERRA, RADARSAT-2 as of May 24, 2010 displaying the polygon that represents anomaly in red.

Figure 6

Results:

Figure 6 is a MODIS image for May 24, 2010 that shows sun-glint reflecting from the oil plume of the Deepwater Horizon disaster. Figure 7 shows the results of processing MODIS data for the same date using SOM algorithm. We can clearly observe that SOM was very successful in picking the regions of oil slick - figure 7 is in very good agreement with figure 5 and 6. We also see how social media data combined with SCAT data on figures 3 and 4 agrees well with the results of SOM in figure 7. We are in the process of developing an approach to programmatically picking the group from SOM that is most likely to represent the oil plume.

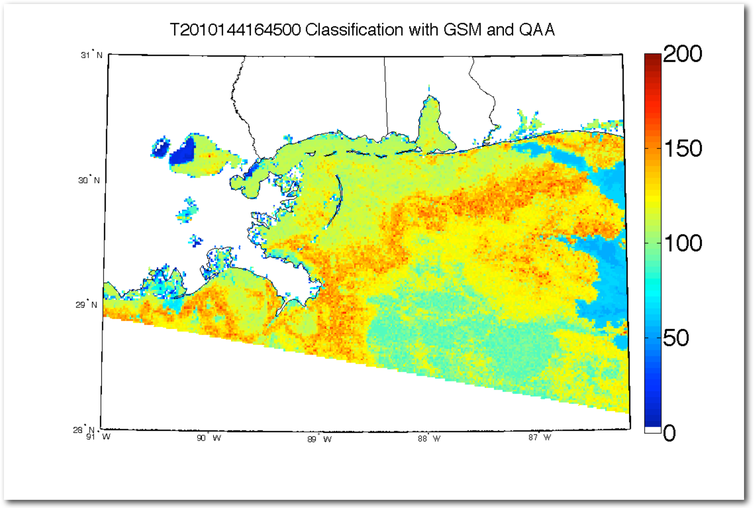

Figure 7

Future Work:

Currently, the input space to the SOM was picked by trial and error until satisfactory map was produced. In the future we plan to develop an approach for selecting the input space automatically. One potential approach that we plan to experiment with is use of genetic algorithms for automated selection of input space.