



Short Communication

Volume 4 Issue 3 - August 2017  
DOI: 10.19080/OFOAJ.2017.04.555638

Oceanogr Fish Open Access J

Copyright © All rights are reserved by Erico N de Souza

# Scaling Up Fishing Activity Visualization with Free Software



Erico N de Souza\*, Johna Latouf and Stan Matwin

*Institute for Big Data Analytics, Dalhousie University, Canada*

**Submission:** July 24, 2017; **Published:** August 23, 2017

**\*Corresponding author:** Erico N de Souza, Institute for Big Data Analytics, Dalhousie University, Canada, Email: [erico.souza@Dal.Ca](mailto:erico.souza@Dal.Ca)

## Short Communication

The Automatic Identification System (AIS) is a protocol imposed by the International Maritime Organization (IMO) for all vessels larger than 300 gross tons or carrying passengers. Generally, organizations use this protocol to identify activities done by vessels during a period of time, but due to the large amount of data, this process needs to be focused on limited periods of time or specific areas. Regular relational databases, such as Postgres, which is a common free relational database to store large data sets, does not keep up with billions of points generated by AIS messages. On the other hand, there is also the issue of visualizing this high number of points in a web interface. This work presents a solution for both problems using only freely available software. The chosen problem is the detection of global fishing activity. Specifically, we want to visualize two types of industrial vessels: trawlers and long liners.

de Souza et al. [1] presented three different approaches to identify fishing activity for trawlers, long liners and purse seiners. Unfortunately, the method for long liner is not as efficient as the other methods, and it also demands knowledge about the vessel type before its execution. In [2] the requirement to identify if a vessel is a trawler or long liner is not necessary, as well, the method reported has significant higher accuracy than the previous approach (the new method reaches 90% accuracy on the fishing detection task on both vessel types, using only close to 60% of long liner data for training). We use this later approach to label each GPS coordinate with fishing or not fishing.

Since the relational databases are not an option, we decided not to use a database, but instead an index, called Apache Solr [3], which has significant differences with a traditional database. The first difference is that it does not guarantee consistency: repeated values may appear. The second difference is the capability to search the data in parallel while automatically configuring and distributing multiple nodes with a partition of the data in each node. The third difference in relation to relational databases is it allows the creation of extra fields dynamically. The creation of dynamic fields allows the connection between the original

AIS data and new computed information from machine learning methods. This means that each point will have aggregated new information that can be used in new applications or to explain specific movement behaviours to an expert. As a way to guarantee that the data we have remains consistent, we also keep a Postgres database with the original data.

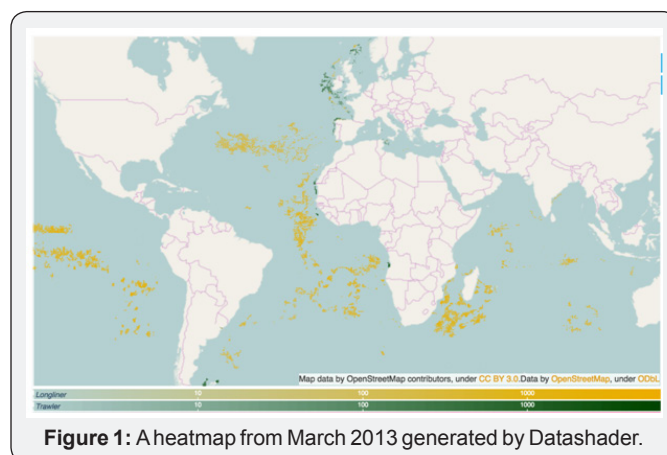


Figure 1: A heatmap from March 2013 generated by Datashader.

Besides the search method, the other aspect to be addressed is how to visualize a large number of points in a web browser. To solve this particular issue, we do not show the raw points directly, but, instead, we compute a heat map that gives a higher level view of the data, and allows the user to interact with it, if it is required to see more details about specific points in the map. The heat map built from the tool can be viewed in Figure 1, and it presents the results of approximately 1,600 vessels during March 2013, where all the vessels listed are trawlers (green) and long liners (yellow). Our fishery visualization tool uses a python package, called Datashader [4], to efficiently build the heat maps with millions of points. The average time to build only the map, i.e. excluding the time to query the data on Solr, is close to 0.5 seconds to render millions of points in a map.

As next steps, we want to expand the tool to visualize other types of fishing vessels. In the short term, we want to add the purse seiner information in these maps. Secondly, the enriched

AIS data will be available to any researcher to download. Currently, downloading the data requires researchers to have an agreement with the Marine Environmental Observation Prediction and Response (MEOPAR) project. We have interest in new applications of this visualization technique to create on demand heat maps of areas with high vessel traffic volume, which will enable researcher to correlate possible impacts for marine life. The tool is currently available at <http://solr.research.cs.dal.ca/fishingobserver/>

### References

1. De Souza EN, Boerder K, Matwin S, Worm B (2016) Improving Fishing Pattern Detection from Satellite AIS Using Data Mining and Machine Learning. PLoS ONE 11(9): e0163760.
2. de Souza EN, Johna L, Xiang J, Matwin S (to appear) An Efficient Algorithm to Predict Fishing Activity.
3. (2017) Solr Development.
4. (2017) Datashader Development Team.



This work is licensed under Creative Commons Attribution 4.0 License  
DOI: [10.19080/OFOAJ.2017.04.555638](https://doi.org/10.19080/OFOAJ.2017.04.555638)

### Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats  
( Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission

<https://juniperpublishers.com/online-submission.php>