pyAIS: Tools for AIS data

Eoin O'Keeffe

August 13, 2015

Abstract

This document outlines how to use the pyAIS library for AIS analysis.

1 Introduction

---

2 Parsing Raw AIS

Ensure the config file is set up correctly. The processing of raw ais occurs in two parts:

- Parse raw ais text files to individual mmsi files. One is created for each vessel. Each files is then stored under a folder whose name is the first two digits of the mmsi of the vessel.
- The second step then converts the mmsi vessels to the imo files.

This process is similar to the processing that took place for the Third IMO GHG report. In addition, it also uses the logging and storing bad data principle that pyrate uses. The data cleaning that takes place is the following:

-

Converting raw AIS data to mmsi files To run the first step is pretty simple:

```
from pyAIS import parser
start_time=time.time()
parser.parse_files_ais_folder(input_type="LR")
print(time.time()-start_time)
```

This parses all the raw data to mmsi files and creates logs and stores any bad data separately. Note that this only uses multithreading (not multiprocessing)

so it doesn't leverage the full power of your machine. Therefore, it may make sense to run a nubmer of these in parallel. You can set one run off point to one data_loc folder and then set another off pointing to another.

Converting mmsi files to imo files If you've just got one output fodler and don't want to run multiprocessing then run the following code:

```
from pyAIS import parser
parser.parse_mmsi_folder_imo()
```

However, if you have multiple output folders or expect processing will take a long time and you can run it on the full machine then use the multiprocessing option. Note that this won't work on ipython.

```
import sys,imp
module_path='C:\\Users\\Eoin\\Documents\\PHD\\Modelling\\Python\\pyAIS'
#module_path='E:\\PHD\\Modelling\\Python\\perform'
if not module_path in sys.path:
    sys.path.append(module_path)
import pyAIS
from pyAIS import parser
parser.parse_mmsi_folder_imo_multiproc()
```

2.1 Parsing to text files on one workstation

2.2 Elastic Map Reduce

A Third party Libraries

- pyrate: A nascent AIS library in python. The parsing techniques are leveraged for this library.

B Config File

```
[file]
data_loc =
wfr_loc=
imo_loc=
ctry_codes_loc=
ports_loc=
tradedownloader_loc=
output_loc=
fixtures_loc=
processing_loc=
[postgres]
```

```
host=localhost
database=
user=
password=
```