Kalpa Publications in Computing
Volume 13, 2019, Pages 123–129
Proceedings of the 1st International Conference on Geospatial Information Sciences

Dictionary–Based Method for Fishing Gear Pattern Detection

Ramón Aranda[1,3] and Hugo Carlos[2,3]

[1] Centro de Investigación Cientifica y de Educación Superior de Ensenada, Unidad de Transferencia Tecnológica (CICESE–UT3); Tepic, Nayarit, México.
[2] Centro de Investigación en Ciencias de Información Geoespacial.
[3] Cátedras del Consejo Nacional de Ciencia y Tecnología (CONACYT), México.
aranda@cicese.mx hcarlos@centrogeo.edu.mx

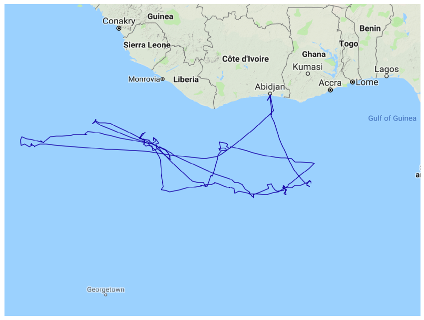
Abstract

Fishing is an ancient practice that dates back to at least the beginning of the Upper Paleolithic period about 40,000 years ago. Nowadays, Fishing is one of the most important activities, as it provides a source of food and economic income worldwide. A key challenge in ecology and conservation is to decrease the Illegal, Unreported and Unregulated fishing (IUU). IUU fishing depletes fish stocks, destroys marine habitats, distorts competition, puts honest fishers at an unfair disadvantage, and weakens coastal communities, particularly in developing countries. One strategy to decrease the IUU fishing is monitoring and detecting the fishing vessel behaviors. Satellite–based Automatic Information Systems (S–AIS) are now commonly installed on most ocean–going vessels and have been proposed as a novel tool to explore the movements of fishing fleets in near real time. In this article, we present a dictionary–based method to classify, by using AIS data, between two fishing gear types: trawl and purse seine. The data was obtained from Global Fishing Watch. Our experiments show that our proposal has a good performance in classifying fishing behaviors, which could help to prevent overexploit and improve the strategies of the fisheries management.

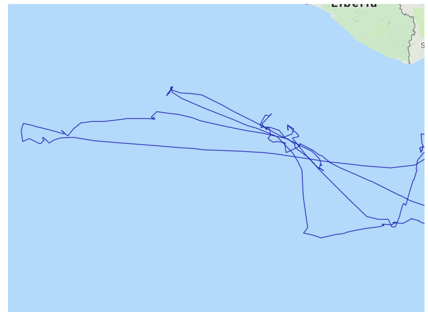
Keywords: AIS, Fishing gear, Dictionary model, Pattern Detection.

1 Introduction

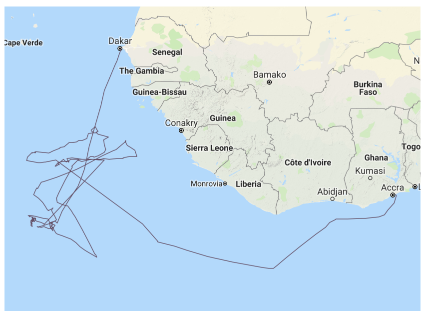
According to the document "Estado Mundial de la Pesca y la Acuicultura, 2018" [5], illicit fishing represents up to 26 million tons of fish per year, more than 15% of the world's annual production. The deficiency of some legal laws, together with the lack of vigilance, lack of implementation of management measures, among others, have hampered the fight against Illegal, Unreported and Unregulated (IUU) fishing. IUU fishing depletes fish stocks, destroys marine habitats, distorts competition, puts honest fishers at an unfair disadvantage, and weakens coastal communities, particularly in developing countries. IUU fishing constitutes a threat to the sustainability of the fishing sector globally, in ecological terms, economic cost, and in



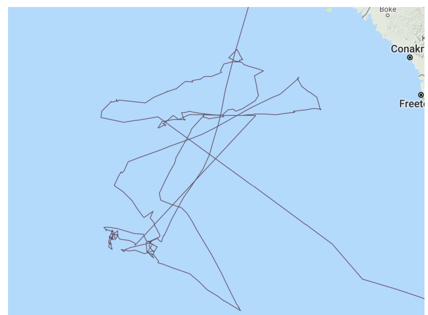
(a)



(b)



(c)



(d)

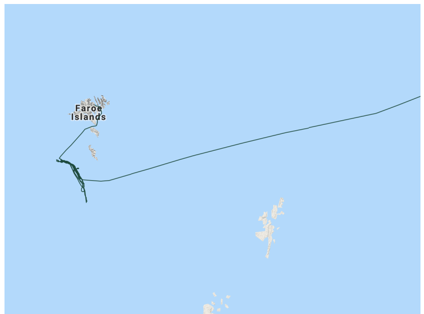
Figure 1: (a) and (c) Lines interpolated representation of raw S–AIS tracks for two individual vessel by using purse seine fishing gear; (b) and (d) more fine–scale of track fishing behavior for a purse seiner of tracks in (a) and (c), respectively..

general of an important source of food and income, as well as the impact on communities that depend on this activity (FAO, 2017). Also, IUU fishing effects directly eco–tourism activities due to it impacts natural habits [6].

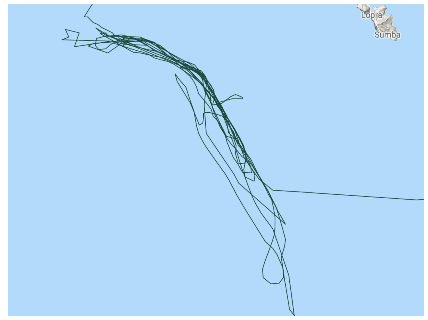
The Sustainable Development Goals (SDG) of the United Nations set a goal for marine fisheries: ”By 2020, effectively regulate fishing exploitation and stop overfishing, illegal fishing, undeclared fishing and unregulated and destructive fishing practices, and apply scientifically based management plans in order to restore fish stocks in the shortest time possible, at least reaching levels that can produce the maximum sustainable yield according to their biological characteristics”. To accomplish this goal, it is necessary many efforts in different areas, one of this is the artificial intelligence to analyze data and see the problem from other perspective.

One strategy to decrease the IUU fishing is monitoring the fishing boats and vessel to determine their behaviors. Satellite–based Automatic Information Systems (S–AIS) are now commonly installed on most ocean–going vessels and have been proposed as a novel tool to explore the movements of fishing fleets in near real time. In contrast, one disadvantage is the poor knowledge about general fishing behaviors of boats and vessels, only there are researches for specific fishing gears [8, 4, 9]

In this article, we present a novel dictionary–based method to classify, by using AIS data, between two fishing gear types: trawl and purse seine. Our method uses a supervised machine



(a)



(b)

(c)

(d)

Figure 2: (a) and (c) Lines interpolated representation of raw S–AIS tracks for two individual vessel by using trawl fishing gear; (b) and (d) more fine–scale of track fishing behavior for a trawler of tracks in (a) and (c), respectively.

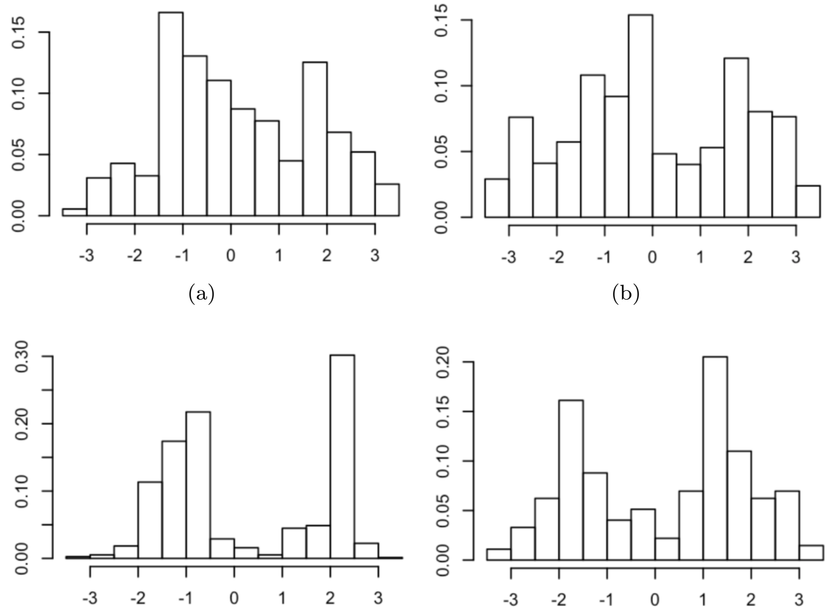
learning strategy that extracts the features of the vessel tracks by analyzing AIS data. Thus, the extracted features are represented by means of normalized histograms. Then, the normalized histograms are used to generate a learned dictionary to represent each fishing gear.

2 Dictionary–based method

Our method uses a supervised machine learning strategy that extracts the features of the vessel tracks by analyzing AIS data. Thus, the extracted features are represented by means of normalized histograms. For this reason, we basically consider a track vessel as the curve described by the vessel when it moves. The sampling of the trajectory implies a step of discretization, i.e., the division of this continuous curve into a number of discrete "steps" connecting successive relocations of the vessel, similar to [13]. Two main classes of trajectories can be distinguished:

- **Trajectories of type I** are characterized by the fact that the time is not precisely know or not taken into account for the relocations of the trajectory;

- **Trajectories of type II** are characterized by the fact that the time is known for each relocation.



(a) (b) (c) (d)

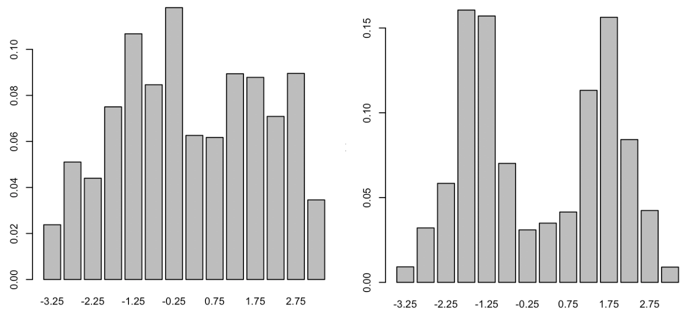
Figure 3: (a) and (b) Histograms of the *absolute angle* for tracks in Figure 1 (purse seine fishing gear); (c) and (d) Histograms of the *absolute angle* for tracks in Figure 2 (trawl fishing gear).

Clearly, track vessels by AIS data information are trajectories type II. A good description of the trajectory can be achieved with the estimation of a minimum set of relatively easily measured parameters.

In our case, we have chosen to characterize all the trajectories by means of the *absolute angle* σ_i between the x direction and the step built by relocations k and $k+1$ [7]. To compute this parameter, we particularly use the implementation in [3]. After computing the *absolute angle*, we compute the normalized histograms of them to characterize a compete track. Then, we can generate supervised learning dictionary to represent each fishing gear as the average of normalized histograms labeled by fishing gear. In our case, we have two normalized histograms in our learned dictionary Φ. Finally, to determine if a track is member of one fishing gear class, we use the Nearest Neighbor [12]. The computation of the distances is made by the Earth Mover's Distance (EMD) [11, 10] between the track and each element of our dictionary. The EMD is explained in the next section.

2.1 Earth Mover's Distance

Give two set of histograms $Q = \{v_i, \alpha_i\}_{i=1}^{N}$ and $J = \{u_j, \beta_j\}_{j=1}^{M}$ where v_i and u_j are the scalar bin values with α_i and β_j their correspond scalar weights of the respective histograms, we denote $d_{ij} = d(v_i, u_j)$ as the Euclidean distance between the bin $i \in Q$ and $j \in J$. Hence, the distances between two histograms $D(Q, J)$ is formulated as the solution of the transportation



(a) (b)

Figure 4: Learned dictionary for different type gears: (a) purse seine fishing gear; and (b) trawl fishing gear.

problem:

$$\begin{aligned}
& D(Q,J) = \min_x \frac{\sum_{i,j} d_{ij} x_{ij}}{\sum_{i,j} x_{ij}} \\
& \text{subject to} \\
& \sum_j x_{ij} \leq \beta_i, \\
& \sum_i x_{ij} \leq \alpha_j \\
& x_{ij} \geq 0 \\
& \sum_{i,j} x_{ij} = \min \left\{ \sum_i \beta_i, \sum_j \alpha_j \right\}.
\end{aligned} \tag{1}$$

In this manner, x_{ij} denotes the transportation flows and represents the amount transported from the $i-th$ supply to the $j-th$ demand. The system of equations 1 is known as the Earth Mover's Distance (EMD). The EMD is a measure which evaluates the dissimilarity between two multi–dimensional distributions in some feature space by using a distance measure between single features. This distance is defined as the minimal cost that must be paid to transform one distributions into the other.

3 Experiments

3.1 Data

For our experiments, we use AIS data available from Global Fishing Watch (GFW)[1]. Although, the raw AIS data are commercial data owned by Orbcomm, GFW has created a sample dataset of labeled anonymized AIS data that can be used to train fishing prediction models [2]. The data can be directly downloaded from https://github.com/GlobalFishingWatch/training-data. For our experiments a total of 20 tracks by each fishing gear were used. Figures 1 and 2 show a example of tracks of raw S–AIS represented by lines for individual vessels and

Table 1: Computed EMD between the dictionary elements and the 20 tracks per each type gear (two main columns). The estimated class is also computed by the Nearest Neighbor (minimum distance). In the table Ps = Purse siene, Tr = Trawl and $\mathbf{Ti}_w$ for $i = 1, 2, \ldots, 20$ with $w \in \{Ps, Tr\}$ indicates the track evaluated for type gear.

Purse Seiner				Trawler			
Tracks	**EMD**$_{Ps}$	**EMD**$_{Tr}$	**Output**	**Tracks**	**EMD**$_{Ps}$	**EMD**$_{Tr}$	**Output**
$\mathbf{T1}_{Ps}$	0.2254637	0.3061086	Ps	$\mathbf{T1}_{Ts}$	0.430946	0.38309	Tr
$\mathbf{T2}_{Ps}$	0.2591721	0.3504665	Ps	$\mathbf{T2}_{Ts}$	0.2863055	0.2347256	Tr
$\mathbf{T3}_{Ps}$	0.2129732	0.4188401	Ps	$\mathbf{T3}_{Ts}$	0.3226546	0.07293751	Ps
$\mathbf{T4}_{Ps}$	0.3005171	0.2975519	Tr	$\mathbf{T4}_{Ts}$	0.7636214	0.5945361	Ps
$\mathbf{T5}_{Ps}$	0.6195521	0.6567899	Ps	$\mathbf{T5}_{Ts}$	0.6281782	0.4930018	Tr
$\mathbf{T6}_{Ps}$	0.4302879	0.4312479	Ps	$\mathbf{T6}_{Ts}$	0.450462	0.2520047	Tr
$\mathbf{T7}_{Ps}$	0.313741	0.223085	Tr	$\mathbf{T7}_{Ts}$	0.4803807	0.2899101	Tr
$\mathbf{T8}_{Ps}$	0.1900463	0.1976318	Ps	$\mathbf{T8}_{Ts}$	0.5631449	0.3464018	Tr
$\mathbf{T9}_{Ps}$	0.1395741	0.3184872	Ps	$\mathbf{T9}_{Ts}$	0.5212029	0.2477887	Tr
$\mathbf{T10}_{Ps}$	0.3957371	0.313737	Tr	$\mathbf{T10}_{Ts}$	0.3563782	0.2606404	Tr
$\mathbf{T11}_{Ps}$	0.4163848	0.4860522	Ps	$\mathbf{T11}_{Ts}$	0.3087636	0.2520101	Tr
$\mathbf{T12}_{Ps}$	0.819418	0.8000955	Tr	$\mathbf{T12}_{Ts}$	0.6371487	0.3471183	Tr
$\mathbf{T13}_{Ps}$	0.2046017	0.3881478	Ps	$\mathbf{T13}_{Ts}$	0.2573653	0.3475851	Ps
$\mathbf{T14}_{Ps}$	0.3226703	0.5555112	Ps	$\mathbf{T14}_{Ts}$	0.2190982	0.2200506	Ps
$\mathbf{T15}_{Ps}$	0.3873934	0.6228465	Ps	$\mathbf{T15}_{Ts}$	0.2365455	0.3537145	Ps
$\mathbf{T16}_{Ps}$	0.1712623	0.3969526	Ps	$\mathbf{T16}_{Ts}$	0.3593299	0.2970012	Tr
$\mathbf{T17}_{Ps}$	0.2203504	0.2824176	Ps	$\mathbf{T17}_{Ts}$	0.4100254	0.1909407	Tr
$\mathbf{T18}_{Ps}$	0.2334088	0.4956399	Ps	$\mathbf{T18}_{Ts}$	0.6263025	0.4075317	Tr
$\mathbf{T19}_{Ps}$	0.2691214	0.4534966	Ps	$\mathbf{T19}_{Ts}$	0.4623226	0.2204181	Tr
$\mathbf{T20}_{Ps}$	0.2216557	0.2705609	Ps	$\mathbf{T20}_{Ts}$	0.2914919	0.2502838	Tr

the two fishing gears. Figure 3 depicts the histograms computed over of the *absolute angle* for tracks in Figures 1 and 2. Note that seeing at a glance the histograms is easily distinguish between fishing gears.

3.2 Results

Before to compute the results, we first must estimate the elements of the dictionary. In our case, we compute each element taking randomly four histograms for each fishing gear type. Figure 4 shows the computed dictionary for the two different type gears: purse seiner and trawler. Table 1 shows the results obtained by proposal method. As we can see, the proposed method has an accuracy of out 80% for the purse seine and 75% for the trawl fishing gear.

4 Conclusions

In this document, we present a novel dictionary-based method to classify two fishing gear types: purse seine and trawl. Our method takes advantages of considering the track vessels as the curve described by the vessel when it moves. The sampling of the trajectory implies a step of discretization, i.e., the division of this continuous curve into a number of discrete "steps" connecting successive relocations of the vessel. This permit us to parametrize the trajectories by means of the *absolute angle* and, thus, to characterize the complete tracks in a normalized histograms. In this way, we can generate a learned dictionary that helps us to classify the tracks depending on their fishing gear. The results of our approach show a good accuracy. Our proposal seems very promising and the results can be improved by parameterizing others features and generating a more complete dictionary.

This research open a new way to analyze track vessels. Our methodology permits to include other fishing gears (or in general vessel behaviors), easily, without the necessity of training again the previous track behaviors. In addition, the track classifying does not require exhaustive computational cost.

References

[1] Global Fishing Watch. https://globalfishingwatch.org. 2019.

[2] Global Fishing Watch. https://globalfishingwatch.org/datasets-and-code/ais-and-other-data/. Accessed on 2019-09-1.

[3] C. Calenge. The package adehabitat for the r software: tool for the analysis of space and habitat use by animals. *Ecological Modelling*, 197:1035, 2006.

[4] Erico N. de Souza, Kristina Boerder, Stan Matwin, and Boris Worm. Improving fishing pattern detection from satellite ais using data mining and machine learning. *PLOS ONE*, 11(7):1–20, 07 2016.

[5] FAO. *EL ESTADO MUNDIAL DE LA PESCA Y LA ACUICULTURA*. Roma, 2018.

[6] Reinaldo Borrego Hernandez Gaspar Gonzalez Sanson. Jorge A. Angulo Valdes, Rodney Borrego Acevedo. Effects of tourism activities on coral reef communities in the punta frances national marine park, cuba. *Rev. Invest. Mar.*, 28(2):159–175, 2007.

[7] L.M. Marsh and R.E. Jones. The form and consequences of random walk movement models. *Journal of Theoretical Biology*, 133(1):113 – 131, 1988.

[8] M. I. Marzuki, R. Garello, R. Fablet, V. Kerbaol, and P. Gaspar. Fishing gear recognition from vms data to identify illegal fishing activities in indonesia. In *OCEANS 2015 - Genova*, pages 1–5, May 2015.

[9] M. I. Marzuki, P. Gaspar, R. Garello, V. Kerbaol, and R. Fablet. Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories. *IEEE Journal of Oceanic Engineering*, 43(3):689–699, July 2018.

[10] Ofir Pele and Michael Werman. Fast and Robust Earth Mover's Distances. In *2009 IEEE 12th International Conference on Computer Vision*, pages 460–467. IEEE, September 2009.

[11] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. The Earth Mover's Distance as a Metric for Image Retrieval. *International Journal of Computer Vision*, 40(2):99–121, November 2000.

[12] B. W. Silverman and M. C. Jones. E. fix and j.l. hodges (1951): An important contribution to nonparametric discriminant analysis and density estimation: Commentary on fix and hodges (1951). *International Statistical Review / Revue Internationale de Statistique*, 57(3):233–238, 1989.

[13] Peter Turchin. *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants. Peter Turchin.* Sinauer Associates, Sunderland, MA., 1998.