COBAC: An Adaptive Transhipment Station Localization for Reducing IUU Fishing Practices

Naman Saxena
Indian Institute of Science
Sakshi Agarwal
Jaypee Institute of Information Technology
Adwitiya Sinha
mailtoadwitiya@gmail.com

Jaypee Institute of Information Technology https://orcid.org/0000-0002-9870-7383

Research Article

Keywords: Transhipment Localization, Wireless Transhipment Station, Machine Learning, Clustering, Illegal Transhipment, IUU Fishing Practices

Posted Date: June 2nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1089527/v1

Version of Record: A version of this preprint was published at Wireless Personal Communications on April 26th, 2024. See the published version at https://doi.org/10.1007/s11277-024-10974-y.

Abstract

The global consumption of fish products is increasing on the scale of millions of tonnes every year. This makes the aquaculture industry as one of the leading sectors to provide food, employment, and ensuring a sustainable livelihood. The implication of rapid growth in global fish production and massive consumption is causing productivity burden on the fisheries management to meet the market demands. This eventually leads to aggravated competition within the fishing networked community. For surviving the competition and increased pressure, few people from fishing community often indulge in various kinds of illegal fishing activities. Illegal, Unreported and Unregulated (IUU) fishing happens to be a major problem plaguing the fish production. Our research proposes a solution based on official transhipment station that solves the problems of illegal transhipment activities, thereby allowing transhipment to continue in a legal and safe manner. We have proposed Cost Optimisation Based Adaptive clustering (COBAC) algorithm that takes into consideration various operational cost and provides the location of establishment of the wirelessly operating transhipment stations in the ocean. The performance of the proposed transhipment was compared with random, greedy and heuristic approaches. Also, the experimentation results show that our proposed COBAC algorithm consumes one-tenth execution time as compared to Brute force clustering and produced result with 0.1% relative error.

1. Introduction

The consumption of fishes has increased substantially all around the globe during the past few decades. The annual population growth rate of the world stood at 1.6 per cent while the annual fish food consumption growth rate is recorded as 3.1 per cent for the period 1961-2017 (FAO, 2020). Per capita fish consumption has increased from 9.0 Kg in 1961 to a record breaking of 20.5 kg in 2018 (FAO, 2020). The increased demand has put enormous pressure on the fishing community. As a consequence of which cases of involvement of fishing boats operating outside the permitted region and resorting to tactics to avoid reporting fish catch has become rampant these days. In the light of increased demand on fish production, Illegal, Unreported and Unregulated (IUU) fishing happens to be one of the major stumbling blocks in the process (Berveridge et al, 2013).

1.1 IUU Fishing & Its Adverse Implications

Illegal(I) fishing refers to the practice of catching fishes in disregard to fishery regulation and violating norms of Exclusive Economic Zones (EEZ)(Camilleri). Unreported(U) fishing is the practice of hiding or misreporting actual amount of fish landing by a fishing vessel. Unregulated (U) fishing refers to the act of operating in the region of a Regional Fisheries Management Organisation (RFMO), to which the country of the concerned fishing vessel does not belong, in non-conformance with the management and conservation norms of RFMO. IUU fishing accounts for 26 million tonnes of fishes being caught, and consequently is responsible for a global loss of 10 billion to 23 billion USD per year (Agnew et al., 2009). There are several other terrible consequences of IUU fishing such as over-exploitation of marine resources and heightened threat to endangered species.

Transhipment refers to the process of transferring the fishes from wirelessly operating fishing vessels to reefer cargo ships. During this phenomena, the fishing vessels often requires refuelling tanks and storing the catches. This allows the fishing vessels to stay longer in the ocean without frequently returning to the home port. Though, the transhipment of fishing vessels seems to be a great strategy to allow economical operation, however, it often gets indulged in hiding and smuggling illegal catches, thus leading to IUU fishing in seas and ocean (Ewell et al., 2017). Sometimes, fishing vessel also trap crew members forcefully on board, thereby violating labour and human rights. The situation is particularly worse in the countries where law enforcement is weak. For instance, incompetent monitoring of waters by countries of West Africa makes it a breeding ground for drugs and weapons smuggling (Bondaroff et al., 2015). The oceans areas susceptible to illegal transhipment activities include Indian Ocean, waters around Southeast Asia (Yea, 2016), Atlantic off West Africa (EJF, 2010) and Western Pacific. Fish species like tuna, Russian pollock, salmon along with crab and wild shrimp are specially threatened by unregulated transhipment (Ewell et al., 2017). Further, not only the monetary aspect of the problem is colossal, but also the disincentive to the small fishermen is huge as the amount of legal catch is cut short by the illegal entrants. If such issues are left untouched or improperly handled they could lead to provoking those who follow rules aptly to indulge into the same. Therefore counteractive measures against IUU fishing is the need of the hour (Azzi et al, 2019, Sunny et al., 2020).

1.2 Related work

A lot of work has been done to tackle IUU fishing. Initially most of the work was in form of policies and laws stipulated by national, international fisheries management bodies. One of the most widely used methods is to decrease the incentive of IUU fishing using trade and market-related policies (Latun et al, 2013; Hosch, 2016). Higher prices are paid for fishes that have proper documentation according to Catch Documentation Scheme of fishery management body. Secondly, fishes caught in the country supporting IUU fishing, are imposed with high tariffs so as to make them unattractive to customers (Le Gallic et al., 2006). Adoption of measure on regional as well as national level along with strict compliance with international regulatory framework could also go long way in reducing the incidents of IUU fishing (Vince et al., 2007; Johns, 2013). In 2009 the United Nations Food and Agriculture Organization (FAO) released Port State Measure Agreement (PSMA), under which the port states have to deny services to fishing vessels involved in IUU fishing (Flothmann et al., 2010). Port measures were effective but suffered from lack of universal implementation. Another way to force the fishing vessels to not involve in IUU fishing is to make amendments to insurance policy. Fishing vessels with history of illegal fishing should be denied liability insurance (Soyer et al., 2018). The problem of IUU fishing could also be seen from the lens of criminology. Rational choice theory and Situation crime prevention theory allow reducing the case of IUU fishing in a systematic and scientific manner. These theories try to find a pattern in the crime and investigate the environmental facts that promoted the crime. Finding patterns in the crime allows coming up with prevention measures. A few broad ways of prevention could be to reduce the access to resource, to increase the risk of being caught (Petrossian, 2015). On the similar lines crime script analysis could be used for response generation (Petrossian et al., 2018). Using crime script analysis, IUU fishing is treated as a crime and it is divided into various stages like Preparation, Entry, Target selection, Doing, Exit. Subsequently various responses in the form of policies are devised corresponding to each stage of the crime.

Game theory also finds its application in reducing IUU fishing. The problem is modelled using Green Security Games (GSGs) approach which provides defender strategy to provide response to attacker (Fang, 2015). Algorithms could be developed for resource allocation and scheduling like patrol boat scheduling. Reinforcement learning is also used to provide solution (Akinbulire, 2017). It is used to model the problem as pursuer-evader game. Using various episodes of training, autonomous agent is able to learn strategies to chase the absconding fishing vessel. Reinforcement learning could be used to provide autonomous patrol boats which could be used in response to an event of illegal fishing. Data is considered the new gold, therefore, information sharing using tools like common database could help the countries cooperate. Regional Fishing Vessel Record database was created so that Association of Southeast Asian Nations (ASEAN) member states could access and provide information regarding fishing vessels that are operating in the EEZ of another country or engaging in IUU fishing (Matsumoto et al., 2012; Saraphaivanich et al, 2016). Since, it is not feasible to comprehensively span across the ocean with patrolling boats for tracking the illegal practices, hence, computational techniques are employed to classify and track the fishing activities through wireless communication.

Classification algorithms are used to track and categorize a particular fishing vessels activity as illegal or legal based on certain oceanographic parameters. These oceanographic variables include sea surface temperature, chlorophyll, and seascape (Woodill et al., 2020). Detection of fishing gear of the vessel could help in deciding whether a fishing vessel is exhibiting unexpected behaviour at sea. Detection of fishing gear also helps in verifying whether a fishing vessel is operating by registering wrong credentials. Such suspicious vessels are likely to be involved in IUU fishing. Vessel Monitoring System (VMS) trajectories are used to predict the fishing gear with the help of classifiers like Support Vector Machine (SVM) and Random Forrest (RF) (Marzuki et al, 2015, 2017). A wireless distributed Automatic Identification System (AIS) is used by vessels for the purpose of navigating and avoiding collision in sea. It can be used to provide vital information about the vessel. Global Fishing Watch used AIS messages to develop a classification model that predict when and where a fishing vessel is engaged in fishing and for how long. This allowed the regulatory agencies to know when a fishing vessel is operating in a prohibited area or if it is involved in overfishing (Merten et al, 2016). Most robust detection of illegal fishing could be accomplished by combing AIS, VMS and Synthetic Aperture Radar (SAR) imagery. Vessels without VMS or with turned off AIS transceiver could be detected using SAR imagery (Longépé et al., 2018). The detection of illegal activity using SAR and Multi Spectral Imager (MSI) data from satellite in conjunction with AIS and VMS has provided a complete solution for surveillance (Kurekin et al., 2019). AIS messages could also be used to detect potential transhipment events at sea (Chuaysi et al., 2020). AIS messages allow us to measure speed of the vessels as well as duration for which the vessel is travelling at that speed. Using certain threshold for speed of the vessel and duration of the event, transhipment events could be identified (Miller et al, 2018).

1.3 Proposed Solution Approach

According to the literature survey, most of the methods are employed to tackle IUU fishing for strengthening the wireless surveillance. A lot of work has been done using AIS and VMS data to detect illegal activity. Transhipment events can be detected using machine learning algorithms on AIS messages. Vessels identified as part of transhipment event could be marked as being involved in illegal activities and face the necessary action from the regulatory authority. But there is a caveat to the above approach. Transhipment is a necessary evil. Transhipment is useful as it allows fishing vessels to make their operation economically efficient and at the same time save time. Therefore, a safe way out of this situation is to establish Transhipment station in high seas and ocean. Transhipment stations will work as the authorized station only where transfer of fishes from one vessel to another vessel is allowed. If the transhipment happens at any other location by the very definition of this approach, such transhipment is illegal. We propose Cost Optimisation-Based Adaptive Clustering (COBAC) technique to find the optimal locations for setting up the wireless transhipment stations by clustering the fishing events.

1.4 Research Highlights

- The global consumption of fish products is increasing on the scale of millions of tonnes annually.
- For surviving aggravated competition, fishing community people often indulge in IUU fishing.
- We have proposed Cost Optimisation-Based Adaptive Clustering (COBAC) to address the issue.
- COBAC considers operational cost to compute the optimal location of wireless transhipment stations in the ocean.
- Our algorithm addresses IUU fishing problems for sustainable fisheries management.

2. Methods & Models

Two obvious questions arise because of this approach: 1)Where will the Transhipment stations be established? 2) Is the operation cost of Transhipment stations feasible? The answer to the above question can be found using a machine learning optimization algorithm that will provide as output the number and location of Transhipment stations. The algorithm will take as input various cost that will be incurred due to this solution. First is the economic cost of setting up stations, and then comes the annual cost of operation, followed by the cost of extra fuel required by fishing vessel to travel to the Transhipment stations. The algorithm provides a complete solution.

2.1 Solution Formulation

The solution involves using a clustering algorithm. The algorithm needs data regarding the fishing activity in seas and ocean. It is provided by Global Fishing Watch's fishing effort data (Kroodsma et al, 2018). The dataset contains information about 2.83 million fishing events in a single year. In order to find location for Transhipment stations, there is a requirement of organising the fishing events into grids based upon the latitude and longitude of fishing event. The coordinates of a grid's centre will be the location of Transhipment station, if that grid is selected for setting up Transhipment station. Further there is a requirement of minimizing cost function $F\left(n_{st}\right)$ along with clustering (Eq. 1).

$$minF\left(n_{st}\right) = C_{ves} \times D_{total} + n_{st} \times C_{st} s.t n_{min} \le n_{st} \le n_{max} (1)$$

C_{ves} is the cost of covering unit distance in ocean by a fishing vessel. D_{total} is the sum of distance travelled by all fishing vessels to reach to transhipment station. C_{st} is the annual cost of operating a transhipment stations. n_{st} is the number of transhipment stations. n_{min} and n_{max} is the minimum and maximum number of transhipment stations allowed respectively. If n_{st} is increased then annual cost of operating transhipment stations ($n_{st} \times C_{st}$) is increased and if n_{st} is decreased, cost of fuel ($C_{ves} \times D_{total}$) increases. Therefore the trade-off between both the costs is non-trivial.

2.2 Proposed COBAC & Other Algorithms

In order to reduce the time complexity of the clustering algorithm the world map is divided into $2° \times 2°$ grids. Consequently 16200 (360/2 × 180/2) grids were generated using Cluster grid generation algorithm (Algorithm 1). Since approximately 70% of earth surface is covered by ocean and not all places on ocean witness fishing activity, we finally get $n_{max} = 2360$ grids for clustering. The $rfmo()$ function takes a grid as input and returns the RFMO that should be responsible for the grid based upon country flag of ships in that grid. The $location()$ function provides the coordinates of the centre of the grid. The $ship\_count()$ function provides the count of fishing ships active in the grid in Algorithm 1. After creation of grids, information regarding the grid centre coordinate is available. Using grid centre location, clusters of grids could be formed. Now the important task is to determine the number of clusters. There are several ways to determine the number of clusters. One way is to try every number of clusters from 1 to n_{max}, which, however, is quite time consuming. Therefore, proposed Cost Optimization-Based Adaptive clustering (COBAC) algorithm (Algorithm 5) was proposed to determine the optimal number of clusters in less number of iterations (Eq. 2-8). For each cluster of grids formed by the algorithm, there is a transhipment station. Here, $n_{st}(t)$ refers to the number of transhipment stations at t^{th} iteration of proposed COBAC algorithm.

Algorithm 1: Cluster Grid Generation
procedure $grid\_generation(data, dim)$
sort $data$ in ascending order according to latitude
divide $data$ into $bins$ of dim degree according to latitude
for each $b \in bins$:
sort b in ascending order according to longitude
divide b into $grids$ of dim degree according to longitude
for each $g \in grids$:
$grids[g]['rfmo'] = rmfo(g)$
$grids[g]['location'] = location(g)$
$grids[g]['ship\_count] = ship\_count(g)$
end for
end for
return $grids$

Table 1
Symbols used in Proposed COBAC algorithm

Symbols	**Definition**
$n_{st}(t)$	Number of transhipment station at t^{th} iteration of proposed COBAC algorithm
$F\left(n_{st}(t)\right)$	Total cost of setting up and operating $n_{st}(t)$ transhipment stations
$\Delta(t)$	Fractional change in number of transhipment station at t^{th} iteration
$\delta(t)$	Fractional change in cost function (Eq. 1) at t^{th} iteration
$\theta(t)$	Ratio of fractional change in cost function and fractional change in number of transhipment station at t^{th} iteration
$\sigma(t)$	Sign of $\theta(t)$ t^{th} iteration
$\pi(t)$	Change in number of transhipment station at t^{th} iteration
$\rho(t)$	Momentum factor for change in number of transhipment stations at t^{th} iteration

$$\Delta(t) = \frac{n_{st}(t) - n_{st}(t-1)}{n_{st}(t-1)} (2)$$

$$\delta(t) = \frac{F\left(n_{st}(t)\right) - F\left(n_{st}(t-1)\right)}{F\left(n_{st}(t-1)\right)} (3)$$

$\theta(t) = \frac{\delta(t)}{\Delta(t)} (4)$ $\sigma(t) = \frac{\theta(t)}{|\theta(t)|} (5)$

$$\rho(t) = \left(\frac{\sigma(t) + \sigma(t-1)}{2} \times \rho(t) + |\sigma(t) - \sigma(t-1)|\right) \times (0.5 + |\theta(t)|) (6)$$

$$\pi(t) = \pi(t-1) - \sigma(t) \times \rho(t) (7)$$

$$n_{st}(t) = \min\left(\max\left(transform(\pi(t)) + n_{st}(t-1), n_{min}\right), n_{max}\right) (8)$$

$$transform(x) = \begin{cases} -1, \ -1 < x < 0 \\ 1, 0 \leq x < 1 \\ floor(\pi(t)), \ otherwise \end{cases} (9)$$

$$floor(x) = \lfloor x \rfloor (10)$$

The COBAC algorithm after the initial iteration will provide a guess for the number of transhipment stations. The only thing preventing us from using K-means is that we did not know the number of clusters. Now that number of clusters is known, we should be able to use K-means. However, usually the clustering algorithms use commutative distance measure (Eq. 11), but because of grids the distance measure is no longer commutative (Eq. 12). Therefore, we need a clustering algorithm that works on a distance measure that follows Eq. 12, to minimize the cumulative cost of clustering, D_{total} (Eq. 13).

$$Distance(u, v) = Distance(v, u) (11)$$

$$Distance(u, v) = \alpha \times Distance(v, u), \ \alpha \in R - \{0\} (12)$$

$$D_{total} = \sum_{i=1}^{n_{grids}} \min_{C} DistEarth\left(center\left(G_i\right), center\left(C_i\right)\right) (13)$$

Here, n_{grids} is the total number of grids with fishing activity. G_i (Fig. 1 blue blocks) refers to the i^{th}grid and C_i (Fig. 1, yellow blocks) refers to the grid that is the centre of the cluster of grids.

Further, $centre(x)$ returns the latitude and longitude of a centre of the grid$x$. $DistEarth(u, v)$ refers to distance in kilometres between two locations, *u* and *v* according to the Geographic Coordinate System

(GCS) (Eqs. 14-17). u_{lat} and u_{lon} represents the latitude and longitude of the location u. Δ_{lat} is the difference between the latitude and Δ_{lon} is the difference between longitude of the two locations, u and v.

$$DistEarth(u, v) = 6373 \times 2 \times atan2\left(\sqrt{a}, \sqrt{1-a}\right)(14)$$

$$a = \sin\left(\frac{\Delta_{lat}}{2}\right)^2 + \cos\left(u_{lat}\right) \times \cos\left(v_{lat}\right) \times \sin\left(\frac{\Delta_{lon}}{2}\right)^2(15)$$

$$\Delta_{lat} = u_{lat} - v_{lat}(16)$$

$$\Delta_{lon} = u_{lon} - v_{lon}(17)$$

The proposed COBAC algorithm is given as Algorithm 5. It takes as input various costs associated with clustering (C_{ves}, C_{st}), data regarding grids, initial estimate of number of transhipment stations and number of iterations to search for optimal number of transhipment stations. The Algorithm 5 outputs optimal number of transhipment station, complete information (in charge RFMO, location) regarding transhipment station and total optimal cost of operation (Eq. 1). The algorithm also returns the transhipment station responsible for a particular grid. The algorithm uses $station\_generation()$ method given as Algorithm 2 to find the best transhipment stations corresponding to the number of transhipment stations allowed. Algorithm 3 also finds transhipment station by selecting first the grids with most fishing vessel as transhipment station in a greedy fashion. Algorithm 4 finds transhipment station by random selection of grids as transhipment station. However, the best performance is achieved with Algorithm 2.

Algorithm 2: Heuristic Transhipment Station Generation
Procedure $station\_generation\left(H_{val}, n_{st}(t), n_{grids}\right)$
stations = { } // list of transhipment stations
non_stations = { }
for $value \in H_{val}$:
if $(value.\ station \notin non\_stations)$:
non_stations = non_stations ∪ {value.station}
end if
if $(non\_stations.\ length = n_{grids} - n_{st}(t))$:
stations = stations ∪ {value.station}
end if
if $(stations.\ length = n_{st}(t))$:
break
end if
end for
return stations

Algorithm 3: Greedy Transhipment Station Generation
Procedure $station\_generation2\left(all\_stations, n_{st}(t)\right)$
all_stations = sort($all\_stations$, *descending*) // ship count-wise sorted
stations = { } // list of transhipment stations
for $value \in all\_stations$:
if (value $\notin stations$) :
stations = stations $\cup$ { value }
end if
if ($stations.length = n_{st}(t)$) :
break
end if
end for
return stations

Algorithm 4: Random Transhipment Station Generation
Procedure $station\_generation3\left(n_{st}(t), n_{grids}\right)$
stations = { } // list of transhipment stations
while ($stations.length \neq n_{st}(t)$) :
index = random(0,n_{grids})
if ($index \notin stations$):
stations = stations $\cup$ { index }
end if
end while
return stations

Algorithm 5: Proposed Cost Optimization-Based Adaptive Clustering (COBAC) Algorithm
Procedure $adaptive\_search\_clustering(C_{ves}, C_{st}, grids, \pi(0), n_{st}(0), n_{iter})$
$t = 0$
station_count = $n_{st}(0)$
for $i \in \{0, 1 \cdots n_{grids} - 1\}$:
for $j \in \{0, 1 \cdots n_{grids} - 1\}$:
distance[i][j] = $DistEarth$(grids[i].location ,grids[j].location)
distance_2[i][j]= distance[i][j]
end for
distance[i] = sort(distance[i], *ascending*)
end for
for $i \in \{0, 1 \cdots n_{grids} - 1\}$:
for $j \in \{1 \cdots n_{grids} - 1\}$:
H_{val}= $H_{val} \cup$ {*count* : distance[i][j] − distance[i][$j - 1$]) × grids[i].ship_count, *station* : i}
end for
H_{val} = sort (H_{val}, *ascending*)// ascending according to *count*
while $(t \neq n_{iter})$:
stations = $station\_generation(H_{val}, n_{st}(t))$
for $i \in \{0, 1 \cdots n_{grids} - 1\}$:
cluster[i] = $\min(distance\_2[i][j \in stations]).\ station$
$D_{total} = D_{total} + \min(distance\_2[i][j \in stations]).\ value$
end for
$F\left(n_{st}(t)\right) = C_{ves} \times D_{total} + n_{st}(t) \times C_{st}$

Algorithm 5: Proposed Cost Optimization-Based Adaptive Clustering (COBAC) Algorithm
if ($F\left(n_{st}(t)\right) < min\_loss$) :
final_cluster = cluster
min_loss =$F\left(n_{st}(t)\right)$
station_count =$n_{st}(t)$
end if
$n_{st}(t+1) = update()$
$t = t+1$
end while
return station_count ,final_cluster, min_loss

The $update()$ function used in the Algorithm 5 uses Eq. 2-8 in sequential manner to get the next value of number of stations. The $update()$ function is inspired from Gradient Descent algorithm used for optimization in machine learning problems. $\theta(t)$ in Eq. 4 calculates the rate of fractional change of cost with respect to fractional change in number of stations. $\rho(t)$ in Eq. 6 is used as a momentum factor. $\pi(t)$ denotes the change in the number of stations,$n_{st}(t)$ at t^{th} iteration. If the value of $\theta(t)$ is either positive or negative for a number of consecutive iteration, $\rho(t)$ is used to have compounded effect on the change in the number of stations,$\pi(t)$. In short, momentum factor $\rho(t)$ is used to accelerate increase or decrease in the number of stations. The final update takes place using Eq. 8. $transform()$ is used to ensure effective change in the number of station is always a non-zero integer. In the expression for $\rho(t)$ in Eq. 6, an arbitrary value of 0.5 is used to prevent $\rho(t)$ from becoming zero and stalling the algorithm.

Algorithm 6: Brute Force Clustering Algorithm
Procedure $brute\_force\_clustering(C_{ves}, C_{st}, grids)$
t = 0
$n_{st}(0) = 1$
station_count = $n_{st}(0)$
// distance, distance_2 matrices are created as in Algorithm 5
// H_{val} list is created in the same way as Algorithm 5
while (t $\neq n_{grids}$) : // Algorithm 6 uses n_{grids} instead of n_{iter} in Algorithm 5
stations = $station\_generation(H_{val}, n_{st}(t))$
// D_{total}, cluster variables are created as in Algorithm 5
$F\left(n_{st}(t)\right) = C_{ves} \times D_{total} + n_{st}(t) \times C_{st}$
if ($F\left(n_{st}(t)\right) < min\_loss$) :
final_cluster = cluster
min_loss = $F\left(n_{st}(t)\right)$
station_count = $n_{st}(t)$
end if
$n_{st}(t+1) = n_{st}(t) + 1$//Algorithm 6 uses naïve update instead of COBAC update
t = t + 1
end while
return station_count, final _cluster, min_loss

The Brute force clustering (Algorithm 6) could also be used instead of proposed COBAC (Algorithm 5). It finds transhipment stations by iterating through all permissible value of number of transhipment stations allowed. Algorithm 6 runs for n_{grids} iterations while proposed COBAC executes for n_{iter} iterations. Here, n_{iter} is a user provided parameter that should not be greater than n_{grids} in order for COBAC to be more efficient than brute force clustering in terms execution time.

2.3 Dataset Description & Collection

One of the most important datasets used in the study is "Daily Fishing Effort at 10th Degree Resolution by MMSI, version 1.0 (2012-2016)" (Dataset A) (Kroodsma et al, 2018)[27]. This dataset provides information regarding daily fishing activity all around the global. It provides the number of fishing hours spent by a fishing vessel identified with Maritime Mobile Service Identity (MMSI) at a particular latitude and longitude. For the purpose of developing clustering algorithm, only data for the year 2012 was taken. "Fishing vessels, version 1.0 (2012-2016)" (Dataset B) (Kroodsma et al, 2018)[27] dataset provided information pertaining to MMSI of fishing vessel and the ISO 3166-1 alpha-3 code of country to which the vessel is registered. In order to get comprehensive data regarding MMSI and the country it is associated with, several other datasets were used. "Identifying Global Patterns of Transhipment Behavior" (Dataset C) (Miller et al., 2018). The Global View of Transhipment: Revised Preliminary Findings (Dataset D) (Kroodsma et al., 2017) also provided MMSI and country flag information. In few of the datasets complete name of the country in place of ISO code of the country corresponding to an MMSI is present. Therefore a dataset from Github repository is downloaded to get the mapping from complete country name to ISO code of the country (Dataset E) (https://github.com/lukes/ISO-3166-Countries-with-Regional-Codes/tree/master/slim-3). The final dataset prepared is called *FishTank* (Fig. 2).

[IMAGE-C:\Workspace\ACDC\ImageHandler\d3

Initially, dataset of fishing activity with all required attributes is generated by adding standard dataset. The final dataset, has the following attributes: latitude, longitude, fishing hours, MMSI, ISO code. ISO code of the country with which MMSI is associated with, will help in deciding which RFMO is responsible for a particular transhipment station. The size of dataset required to get the location of transhipment station such that Eq. (1) is minimized, is huge.

Fig 3: Flow diagram of proposed COBAC algorithm

Therefore, there was a requirement to divide the entire world map into grids of dimension$2^\circ \times 2^\circ$. For this purpose, Cluster grid generation algorithm (Algorithm 1) was used that the dataset $FishTank$ as input with dimension, $dim = 2$ (fig. 3). The cluster grid generation algorithm returned $2^\circ \times 2^\circ$ grids and information regarding the grids. Information regarding the grids includes the RFMO responsible for the grid, total number of unique MMSI (fishing vessels) present in the grid and coordinates of the centre of the grid. Fig. 5 shows the area of jurisdiction for a limited set of RFMOs for the sake of clarity. Fig. 4 plots the $2^\circ \times 2^\circ$ grids and intensity of colour denotes the number of fishing vessels present. Further proposed COBAC algorithm (Algorithm 5) is invoked. The input parameters include, Distance cost (C_{ves}), Annual operation cost (C_{st}), Information regarding grids, initial change in number of transhipment station ($\pi(0)$), initial number of transhipment station($n_{st}(0)$)and number of iterations (n_{iter}) to execute our proposed COBAC algorithm. The proposed COBAC algorithm (Algorithm 5) returns as output the minimized cost of implementing well-localized transhipment station strategy. It also returns the optimal number of transhipment Fig 4: Distribution of $2^\circ \times 2^\circ$ grids & density of fishing vesselsstations and assignment of all the $2^\circ \times 2^\circ$ grids to a transhipment station.

PICESICESIATTCICCATCCAMLRCACFishIOTCFig 5: Distribution of RFMOs on world map

3. Experimental Results

The proposed COBAC algorithm performs well in terms of time complexity when compared with other clustering algorithm (section 3.1). Its performance is also compared with the Brute force clustering algorithm. In order to arrive at the best version of proposed COBAC algorithm, several version of wireless transhipment station generation algorithms are also explored.

3.1 Time Complexity Analysis

There are various clustering algorithms available which can be used to cluster the fishing events. Disadvantage with a lot of clustering algorithms is their large time complexity (Table 1) (Xu et al., 2015)[28]. Algorithms with large time complexity take a lot of time to generate clusters.

Table 1
Clustering algorithms and their time complexity

Algorithm	Time Complexity	Comment
K-medoids	$O\left(k(n-k)^2\right)$	k is the number of clusters
GMM	$O\left(n^2kt\right)$	k is the number of clusters, t is number of iterations
DBSCAN	$O(nlogn)$	n is the size of input points
CLARANS	$O\left(n^2\right)$	n is the size of input points

Fortunately there are some clustering algorithms with low time complexities as illustrated in Table 2 (Xu et al., 2015)[28]. However, there is a significant shortcoming even with algorithms with low time complexities. These algorithms either take as input the number of cluster or find the number of clusters without minimizing objective function (Eq. 1).

Table 2
Clustering algorithms with low time complexity

Algorithm	Time Complexity	Comment
K-means	$O(knt)$	k is the number of clusters, t is the number of iterations
BIRCH	$O(n)$	n is the size of input points
Wavecluster	$O(n)$	n is the size of input points

Therefore, there was a requirement for a clustering algorithm with low time complexity as well as one that can minimize the objective function (Eq. 1). The time complexity of the most efficient clustering algorithm available till now is $O\left(n_{ac}\right)$, where n_{ac} is the total count of recorded fishing events by vessels.

Our procedure has a time complexity of $O\left(\left(n_{grids}\right)^2 \times n_{iter}\right)$. The maximum value of n_{iter} could be n_{grids}. Therefore effective time complexity is $O\left(\left(n_{grids}\right)^3\right)$. n_{grids} is a constant value (Eq. 18). Hence, the algorithm works in $O(1)$ time complexity (Eq. 19) and at the same time optimizes objective function (Eq. 1).

$$n_{grids} \approx \left(\frac{360}{dim}\right) \times \left(\frac{180}{dim}\right), dim \in \{1,2\cdots 180\} (18)$$

$$O\left(\left(n_{grids}\right)^2 \times n_{iter}\right) \rightarrow O\left(\left(n_{grids}\right)^3\right) \rightarrow O(1) (19)$$

3.2 Performance of Proposed COBAC Algorithm

Proposed COBAC algorithm minimizes Eq. 1 without iterating for all the values of $n_{iter}\left(n_{iter} \in \left\{1,2\cdots n_{grids}\right\}\right)$, where $n_{min} = 10$ and $n_{max} = 2360$. Table 3 shows the performance of COBAC algorithm (Algorithm 5) versus the Brute force clustering algorithm (Algorithm 6). Brute force clustering also provides equally efficient output as COBAC but takes 10 times the execution time of COBAC (Table 3). Our proposed COBAC (Algorithm 5) produces results with a relative error of 0.1% compared to Brute force clustering (Algorithm 6) and the execution time of Algorithm 5 is one-tenth of execution time of Algorithm 6. The total cost obtained in various iterations by COBAC and Brute force clustering is given in Fig. 6 and Fig. 7 respectively. Fig. 8 explicitly shows the different number of transhipment stations explored for obtaining total cost in various iterations by COBAC. Our proposed COBAC algorithm takes 23 iterations to first come across the best solution, while the existing approach takes 425 iterations to attain the best solution.

Table 3
Performance of Proposed COBAC & Brute Force Search Clustering Algorithm

Method	Parameters		Performance Evaluation	
			Execution Time(seconds)	**Min Loss (USD)**
COBAC (Algorithm 5)	$C_{ves} = 200$ $C_{st} = 1000000$	$\pi(0) = 5$ $n_{st}(0) = 200$ $n_{iter} = 200$	248	1490379796.44102
COBAC (Algorithm 5)		$\pi(0) = 7$ $n_{st}(0) = 250$ $n_{iter} = 200$	255	1492878960.10300
Brute force search clustering (Algorithm 6)			2784	1489871291.54837
COBAC (Algorithm 5)	$C_{ves} = 100$ $C_{st} = 5000000$	$\pi(0) = 10$ $n_{st}(0) = 250$ $n_{iter} = 100$	78	2889928747.02890
COBAC (Algorithm 5)		$\pi(0) = 5$ $n_{st}(0) = 150$ $n_{iter} = 200$	131	2902123802.58631
Brute force search clustering (Algorithm 6)			2828	2886381606.48139

Fig 6: Optimal total cost of proposed COBAC

Fig 7: Optimal cost and number of transhipment stations of Brute-force search clustering

Fig 8: Optimal number of transhipment stations of proposed COBAC

Proposed COBAC algorithm is tested with different versions of transhipment station generation algorithms (Algorithm 2, 3, 4). Performance based on various parameters is given in Table 4. Algorithm 2 outperforms all the other algorithms (Algorithm 3, 4) used to get the best transhipment stations corresponding to a particular value of $n_{st}(t)$ (Table 4). Fig. 9 shows the final distribution of transhipment station along with the $2° \times 2°$ grids on map of world, where blue dots represent transhipment stations and feeble reds dots represents grids.

Table 4
Performance of different station generation algorithms

Method	Parameters	Performance	
		Execution Time(seconds)	**Min Loss (USD)**
Heuristic station generation (Algorithm 2)	$C_{ves} = 200$, $C_{st} = 5000000$, $\pi(0) = 10$, $n_{st}(0) = 100$, $n_{iter} = 100$	92	3972562272.54781
Greedy station generation (Algorithm 3)		93	4742851566.48398
Random station generation (Algorithm 4)		80	5071757330.48577
Heuristic station generation (Algorithm 2)	$C_{ves} = 200$, $C_{st} = 5000000$, $\pi(0) = 5$, $n_{st}(0) = 100$, $n_{iter} = 200$	143	3970113641.90638
Greedy station generation (Algorithm 3)		167	4736114266.7288
Random station generation (Algorithm 4)		172	5177855664.14666

4. Discussions

Our research attempts to address the issue of IUU fishing by focusing on the problem of illicit usage of transhipment in oceans. Though, transhipment solves several problems encountered by fishing vessels functioning in the oceans and seas, however falls under the grasp of illegal activities. Our study offers a solution by proposing to organize the transhipment activity being carried out in oceans. Official wireless transhipment stations could be set up at appropriate locations in oceans and will be responsible to arrange interaction between cargo vessels and fishing vessels. The interaction will be monitored by appropriate regulatory authority and consequently this will put a check on all the illegal activities being carried out in lieu of necessary activities. This approach will also make it easier to detect illegal transhipment in oceans. Several methods have already been developed to detect transhipment activity in ocean. Since, the transhipments will only be allowed on official stations, all the rest of the transhipment activity apart from that on official station, could easily be termed as illegal. Automatically, there is a requirement of a system that could provide the location for establishing official transhipment station, so that their operation is economically efficient.

5. Conclusion

Our proposed COBAC algorithm provided by this study delivers the locations of the wireless transhipment stations keeping in consideration the cost of operation of stations as well as the extra cost incurred by fishing vessels to reach the stations. The stations could be managed by appropriate RFMO. Therefore, the algorithm also assigns a RFMO to a transhipment station depending upon the location of transhipment station and country flag of active fishing vessels. COBAC algorithm takes one-tenth execution time as compared to Brute force clustering algorithm and produces result with 0.1% relative error. Our COBAC algorithm was capable to locate optimal number of stations in one-eighteenth number of iterations as compared to Brute force clustering. The algorithm also enjoys a time complexity of $O(1)$ because of using grid structure. For our experimentation, fishing activity of only year 2012 was considered. The algorithm could be made more efficient by using fishing activity data for more than one year. Further, information regarding the routes followed by cargo vessels could also be incorporated to make the whole arrangement more economic for cargo vessels as well. Moreover, the data regarding anchorage points could also be used for deciding whether station should be set up in ocean-bed or on-land.

Declarations

*Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript

*Conflicts of interest/Competing interests: The authors have no relevant financial or non-financial interests to disclose.

*Availability of data and material: The datasets generated and analysed during the current study are available in the FishTank dataset repository at github.com/namansaxena9/FishTank-Dataset

*Code availability: The datasets generated and analysed during the current study are available in the FishTank dataset repository at github.com/namansaxena9/FishTank-Code

*Authors' contributions: All authors contributed to the research base, conception and design. Material preparation, data collection and analysis were performed by Naman Saxena, Sakshi Agarwal and Adwitiya Sinha. The first draft of the manuscript was written by Naman Saxena, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

Akinbulire, T., Schwartz, H., Falcon, R., & Abielmona, R. (2017, November). A reinforcement learning approach to tackle illegal, unreported and unregulated fishing. In 2017 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 1-8). IEEE.

Agnew, D. J., Pearce, J., Pramod, G., Peatman, T., Watson, R., Beddington, J. R., & Pitcher, T. J. (2009). Estimating the worldwide extent of illegal fishing. *PloS one*, *4*(2), e4570.

Azzi, R., Chamoun, R. K., & Sokhn, M. (2019). The power of a blockchain-based supply chain. Computers & industrial engineering, 135, 582-592.

Berveridge, M., Dieckmann, U., Fock, H. O., Froese, R., Keller, M., Löwenberg, U., Merino, G., Möllmann, C., Pauly, D., Prein, M., Quaas, M., Schmidt, J., Schulz, C., Voss, R. and Zimmermann, C. and Maribus, Future Ocean - Cluster of Excellence, IOI International Ocean Institute.(2013). World Ocean Review 2013: Living with the ocean: 2. The Future of Fish-The Fisheries of the Future.

Bondaroff, T. N. P., Van Der Werf, W., & Reitano, T. (2015).Case Study: IUU Fishing in West Africa.The illegal fishing and organized crime nexus: Illegal fishing as transnational organized crime. The Global Initiative Against Transnational Organized Crime and The Black Fish, 48-49.

Camilleri, M. (n.d.). *Illegal, Unreported and Unregulated (IUU) fishing*. What is IUU fishing? | Illegal, Unreported and Unregulated (IUU) fishing | Food and Agriculture Organization of the United Nations. Retrieved November 8, 2021, from fao.org/iuu-fishing/background/what-is-iuu-fishing/.

Chuaysi, B., & Kiattisin, S. (2020). Fishing vessels behavior identification for combating iuu fishing: Enable traceability at sea. *Wireless Personal Communications*, 1-23.

Duncalfe, L. (2019). ISO-3166 Country and Dependent Territories Lists with UN Regional Codes. GitHub. https://github.com/lukes/ISO-3166-Countries-with-Regional-Codes

EJF. (2010). All at Sea–The Abuse of Human Rights Aboard Illegal Fishing Vessels.

Ewell, C., Cullis-Suzuki, S., Ediger, M., Hocevar, J., Miller, D., & Jacquet, J. (2017). Potential ecological and social benefits of a moratorium on transshipment on the high seas. *Marine Policy*, *81*, 293-300.

Fang, F., Stone, P., & Tambe, M. (2015, July). When Security Games Go Green: Designing Defender Strategies to Prevent Poaching and Illegal Fishing. In IJCAI (pp. 2589-2595).

FAO. 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in action. Rome.https://doi.org/10.4060/ca9229en

Flothmann, S., von Kistowski, K., Dolan, E., Lee, E., Meere, F., & Album, G. (2010). Closing loopholes: getting illegal fishing under control. Science, 328(5983), 1235-1236.

Food And Agricultural Organisation of United Nations. 2020. Illegal, Unreported and Unregulated (IUU) fishing. Article available at: fao.org/iuu-fishing/background/what-is-iuu-fishing (accessed 13 August 2020)

Food And Agricultural Organisation of United Nations. Illegal, Unreported and Unregulated (IUU) fishing. Article available at: fao.org/iuu-fishing (accessed 13 August 2020)

Hosch, G. (2016). Trade measures to combat IUU fishing: Comparative analysis of unilateral and multilateral approaches. Geneva, Switzerland: International Centre for Trade and Sustainable Development (ICTSD).

Human trafficking (2020). Stop Illegal Fishing. Article available at: https://stopillegalfishing.com/issues/human-trafficking, (retrieved August 14, 2020)

Johns, M. (2013). Enhancing responsible fishing practices in South East Asia to combat Illegal, Unreported and Unregulated (IUU) fishing. Australian Journal of Maritime & Ocean Affairs, 5(3), 112-119.

Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., Wilson, A., Bergman, B., White, T. D., Block, B. A., Woods, P., Sullivan, B., Costello, C., Worm,B. (2018). Tracking the global footprint of fisheries. Science, 359(6378), 904-908.

Kroodsma, D.A., N.A. Miller, and A. Roan 2017. "The Global View of Transshipment:Revised Preliminary Findings." Global Fishing Watch and SkyTruth, July 2017. Available online at http://globalfishingwatch.org.

Kurekin, A. A., Loveday, B. R., Clements, O., Quartly, G. D., Miller, P. I., Wiafe, G., & Adu Agyekum, K. (2019). Operational monitoring of illegal fishing in Ghana through exploitation of satellite earth observation and AIS data. Remote Sensing, 11(3), 293.

Latun, A. R., bin Ahmad, A. T., Ali, M., Katoh, M., & Sulit, V. T. (2013). Developing market measures to control IUU fishing in Southeast Asia. Secretariat, Southeast Asian Fisheries Development Center.

Le Gallic, B., & Cox, A. (2006). An economic analysis of illegal, unreported and unregulated (IUU) fishing: Key drivers and possible solutions. Marine Policy, 30(6), 689-695.

Longépé, N., Hajduch, G., Ardianto, R., de Joux, R., Nhunfat, B., Marzuki, M. I.,Fablet, R., Hermawan, I., Germain,O., Subki,B.A., Farhan, R, Muttaqin, A.D., Gaspar, P. (2018). Completing fishing monitoring with spaceborne Vessel Detection System (VDS) and Automatic Identification System (AIS) to assess illegal fishing in Indonesia. Marine pollution bulletin, 131, 33-39.

Marzuki, M. I., Garello, R., Fablet, R., Kerbaol, V., & Gaspar, P. (2015, May). Fishing gear recognition from VMS data to identify illegal fishing activities in Indonesia. In Oceans 2015-Genova (pp. 1-5). IEEE.

Marzuki, M. I., Gaspar, P., Garello, R., Kerbaol, V., & Fablet, R. (2017). Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories. IEEE Journal of Oceanic Engineering, 43(3), 689-699.

Matsumoto, K., Chokesanguan, B., Sulit, V. T., & Saraphaivanich, K. (2012). Development of regional fishing vessels record as tool to combat IUU fishing in Southeast Asia. Secretariat, Southeast Asian Fisheries Development Center.

Merten, W., Reyer, A., Savitz, J., Amos, J., Woods, P., & Sullivan, B. (2016). Global Fishing Watch: Bringing transparency to global commercial fisheries. arXiv preprint arXiv:1609.08756.

Miller, N. A., Roan, A., Hochberg, T., Amos, J., & Kroodsma, D. A. (2018). Identifying global patterns of transshipment behavior. Frontiers in Marine Science, 5, 240.

Petrossian, G. A. (2015). Preventing illegal, unreported and unregulated (IUU) fishing: A situational approach. Biological Conservation, 189, 39-48.

Petrossian, G. A., & Pezzella, F. S. (2018). IUU fishing and seafood fraud: Using crime script analysis to inform intervention. The ANNALS of the American Academy of Political and Social Science, 679(1), 121-139.

Saraphaivanich, K., Suthipol, Y., Imsamrarn, N., Chokesanguan, B., & Siriraksophon, S. (2016). Regional fishing vessels record: A management tool for combating IUU fishing in Southeast Asia. Fish for the People, 14(2), 12-17.

Soyer, B., Leloudas, G., & Miller, D. (2018). Tackling IUU fishing: developing a holistic legal response. Transnational Environmental Law, 7(1), 139-163.

Sunny, J., Undralla, N., & Pillai, V. M. (2020). Supply chain transparency through blockchain-based traceability: An overview with demonstration. Computers & Industrial Engineering, 106895.

Vince, J. (2007). Policy responses to IUU fishing in Northern Australian waters. Ocean & coastal management, 50(8), 683-698.

Woodill, A. J., Kavanaugh, M., Harte, M., & Watson, J. R. (2020). Predicting Illegal Fishing on the Patagonia Shelf from Oceanographic Seascapes. arXiv preprint arXiv:2007.05470.

Xu, D., & Tian, Y. (2015). A comprehensive survey of clustering algorithms. Annals of Data Science, 2(2), 165-193.

Yea, S. (2016). Diluted Justice: Protection and Redress for Trafficked Fishermen in Asia. *Transient Workers Count Too*.

Figures

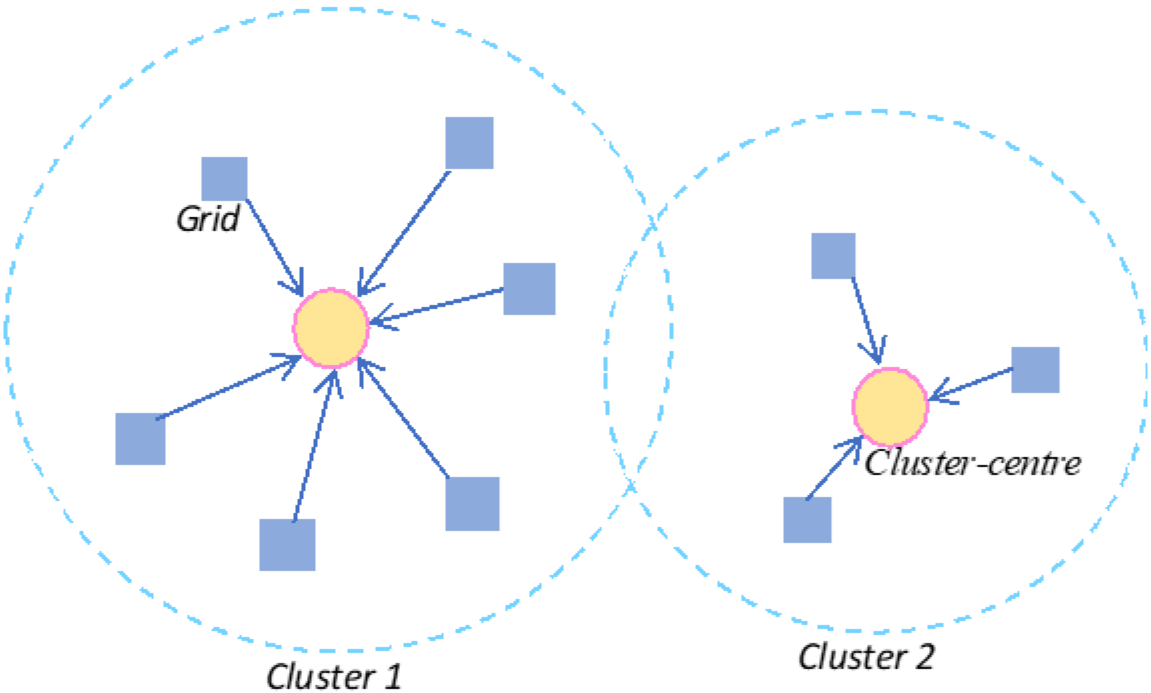

Figure 1

Organisation of clusters and grids

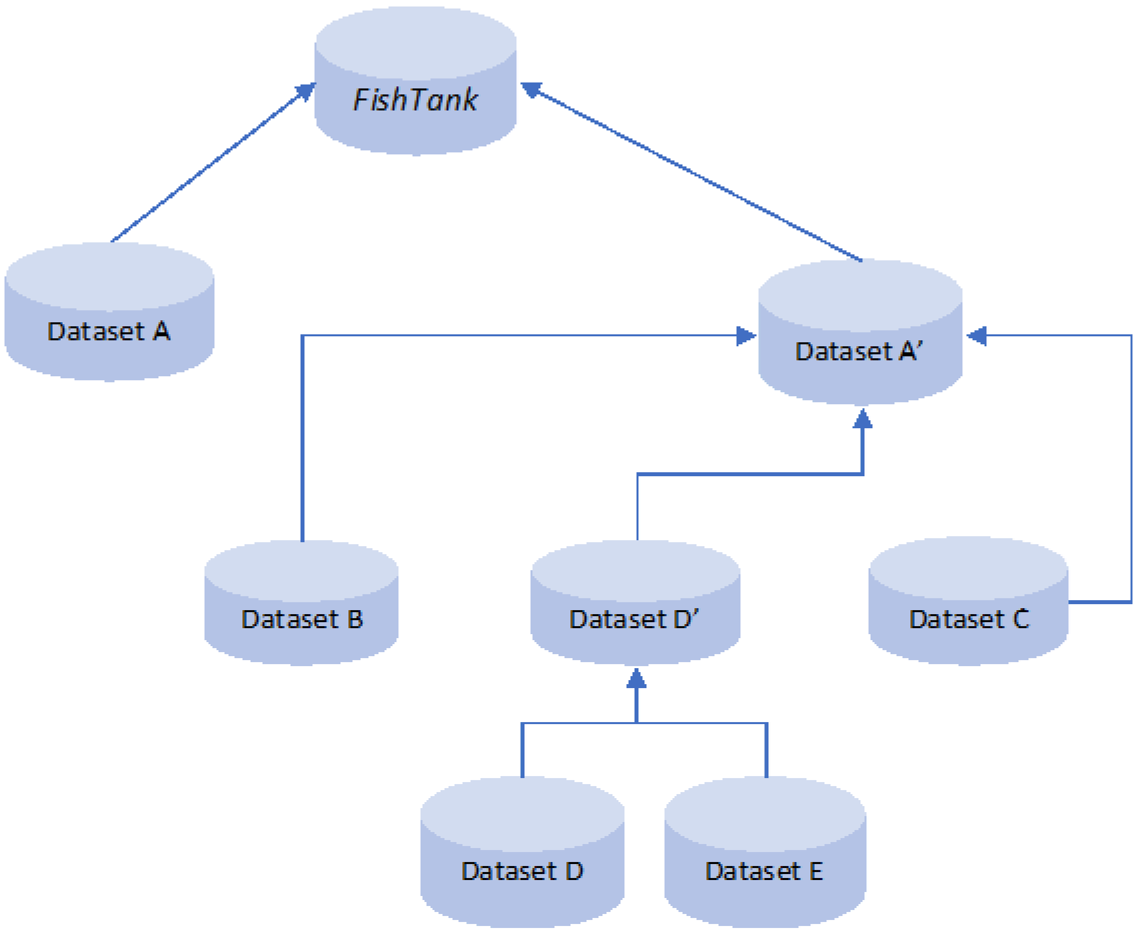

Figure 2

FishTank dataset construction scheme

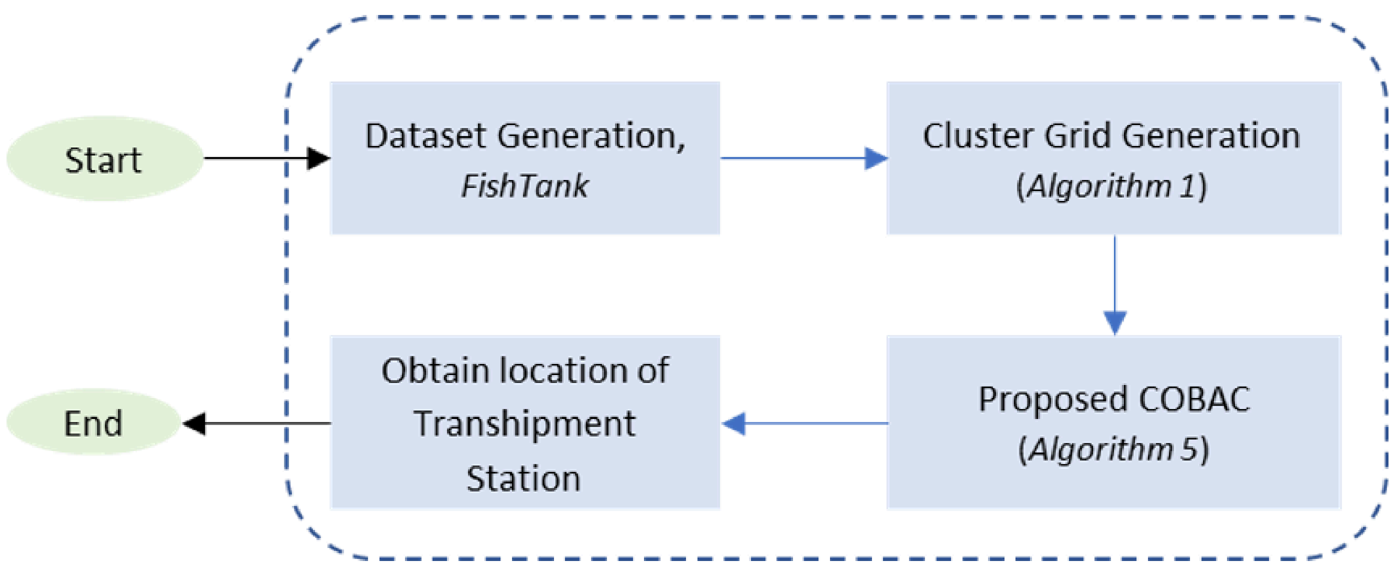

Figure 3

Flow diagram of proposed COBAC algorithm

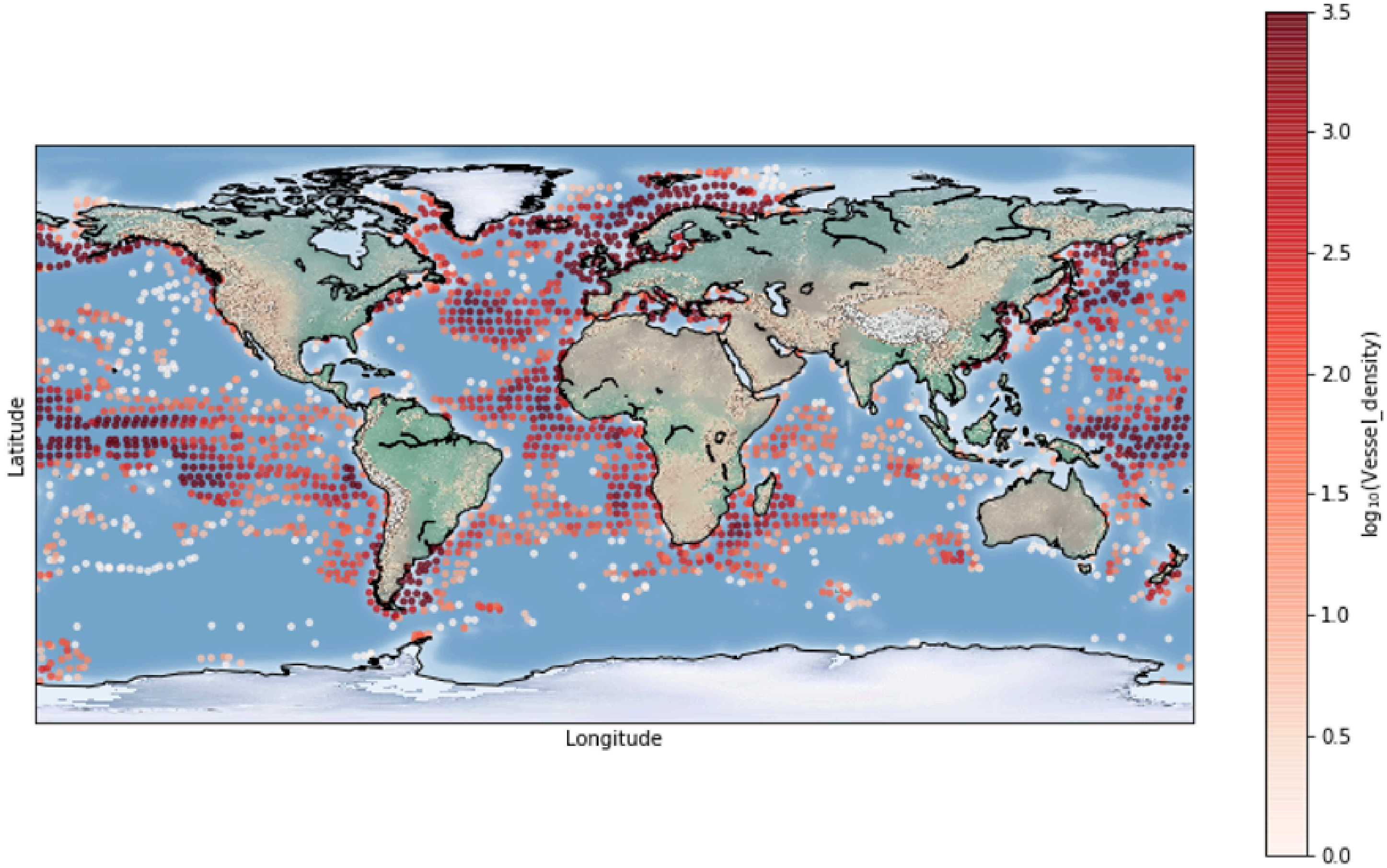

Figure 4

Distribution of grids & density of fishing vessels

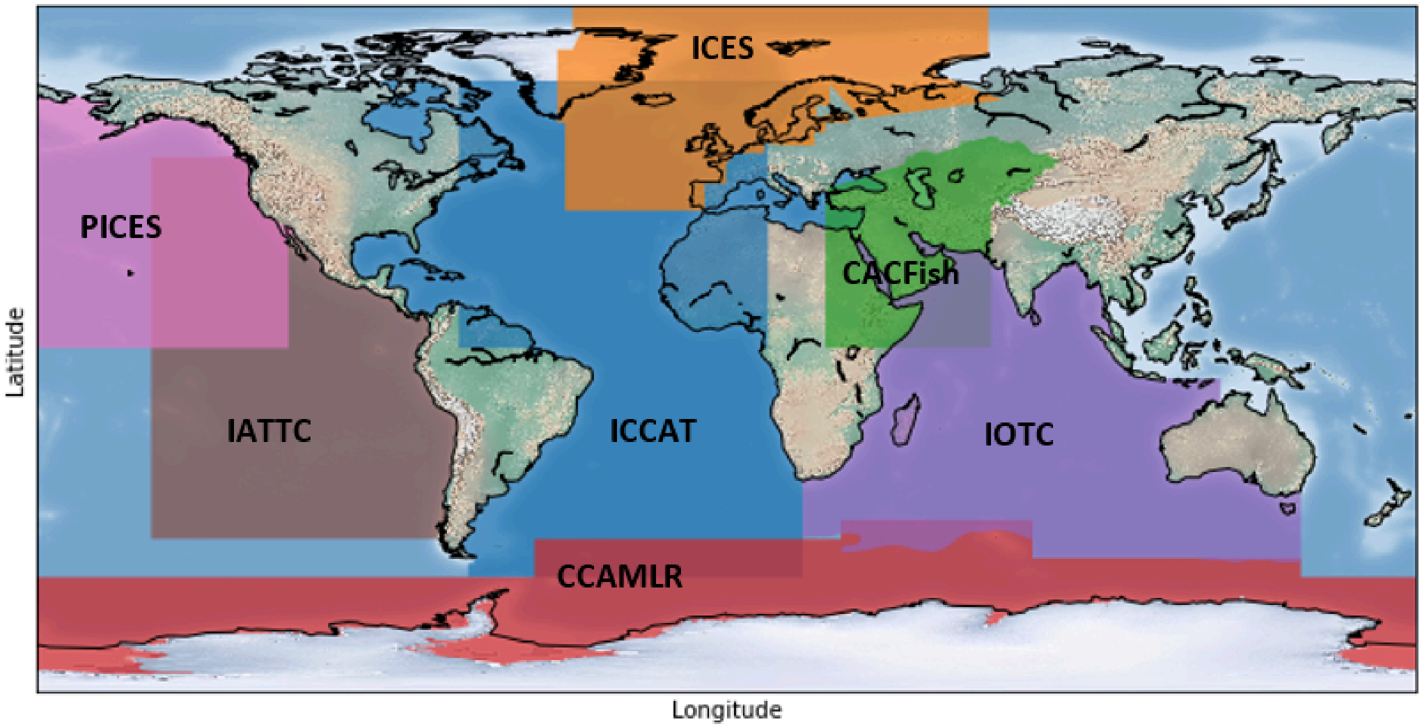

Figure 5

Distribution of RFMOs on world map

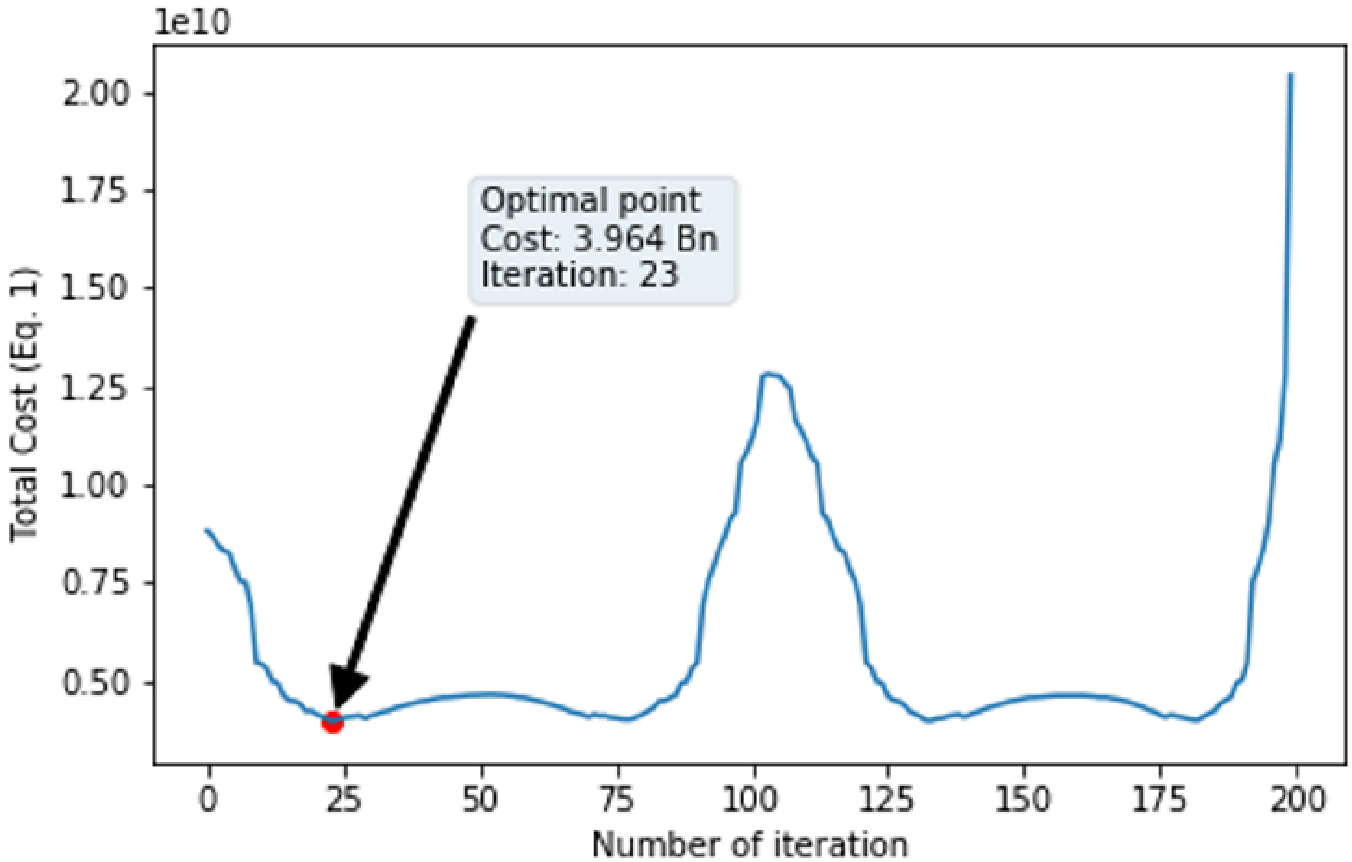

Figure 6

Optimal total cost of proposed COBAC

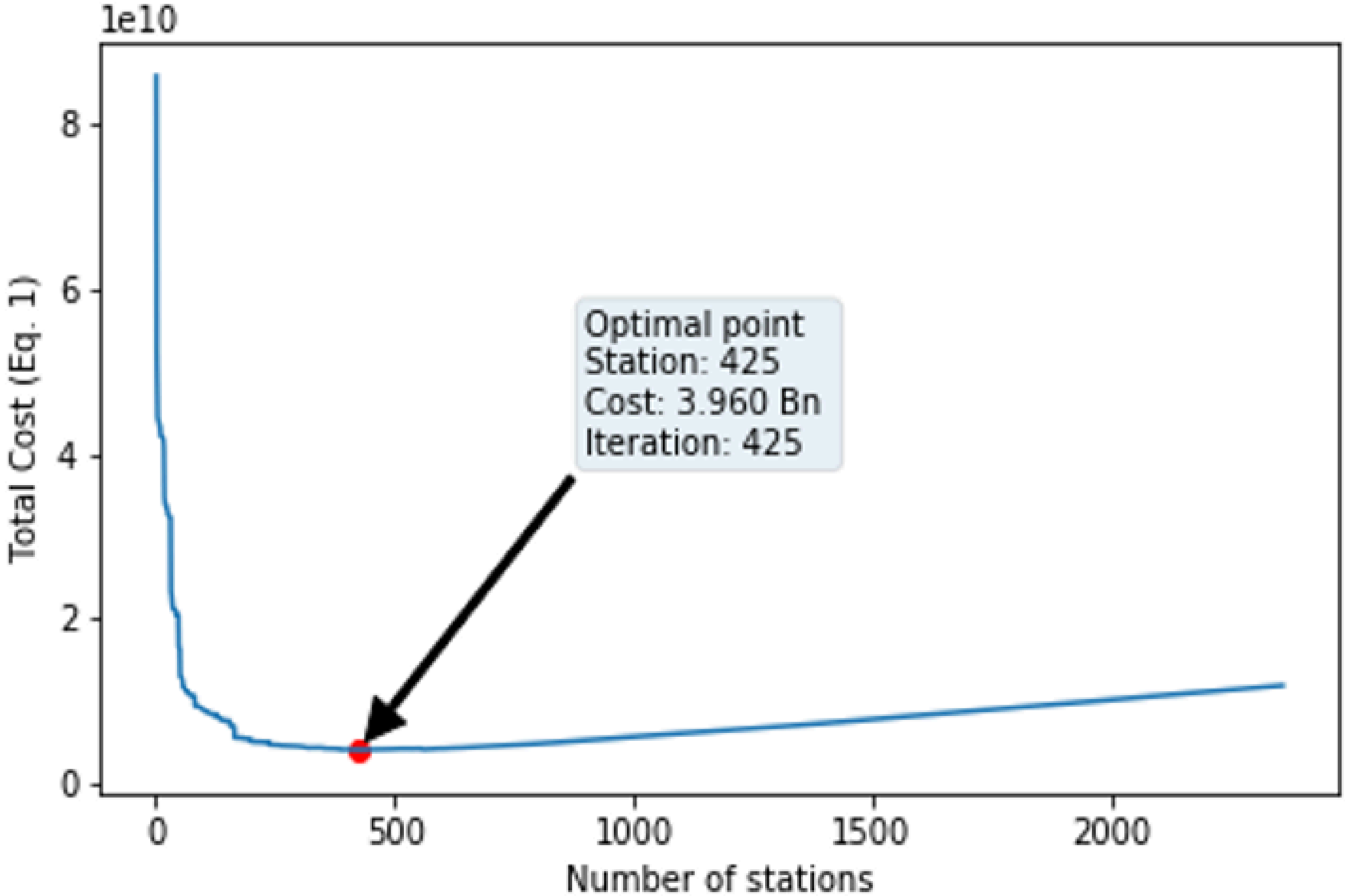

Figure 7

Optimal cost and number of transhipment stations of Brute-force search clustering

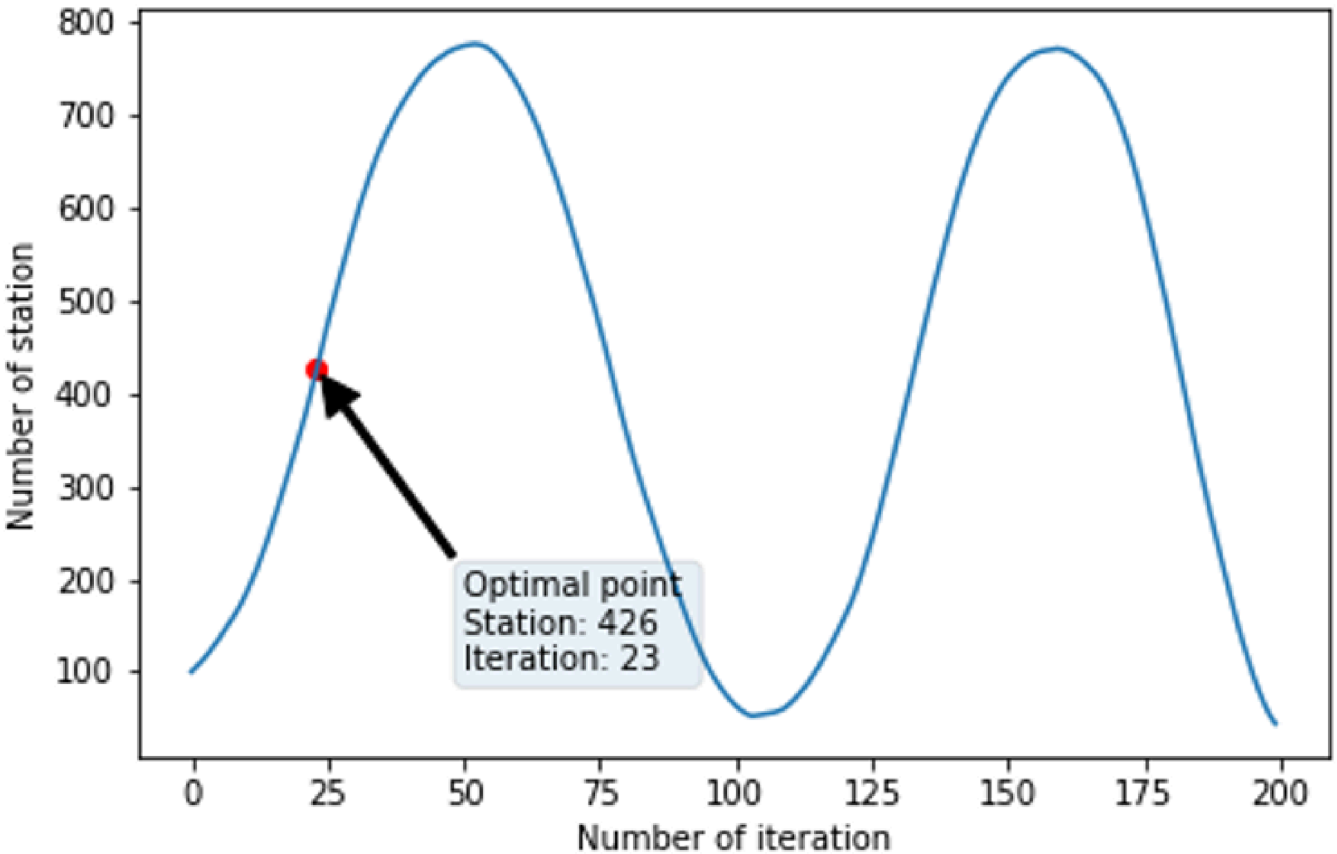

Figure 8

Optimal number of transhipment stations of proposed COBAC