28º International Congress on Waterborne Transportation, Shipbuilding and Offshore Constructions

Rio de Janeiro/RJ, 27th to 29th October 2020

A Dynamic Port Congestion Indicator – A Case Study of the Port of Rio de Janeiro

Vinícius Barreto Martins, UFRJ/COPPE, Rio de Janeiro/Brasil, v1nybarreto@oceanica.ufrj.br
Ramiro Fernandes Ramos, UFRJ/COPPE, Rio de Janeiro/Brasil, ramirofr@oceanica.ufrj.br
Maricruz Aurelia Fun Sang Cepeda, UFRJ/COPPE, Rio de Janeiro/Brasil, maricruzcepeda@oceanica.ufrj.br
Jean-David Caprace, UFRJ/COPPE, Rio de Janeiro/Brasil, jdcaprace@oceanica.ufrj.br

Abstract
Maritime trade plays a key role in the global economy and recent technological developments have accelerated maritime logistics. However, this increase in maritime trade has had an impact on port performance, leading to port congestion in some regions. Few researches employing AIS data has explored the marine traffic congestion, hence the development of a system that makes metrics on ports more accessible is needed. This work employs an innovative methodology to analyze the port congestion level on the port of Rio de Janeiro. From the Automatic Identification System (AIS) data, three algorithms were used to find the convex hull area, the geolocation area, and the average vessel proximity. These algorithms were used to calculate the Port Congestion Indicators (PCIs): i) Spatial Concentration; ii) Spatial Density; iii) Average Service Time. Then, Machine Learning techniques were employed to cluster these indicators into low, medium, and high congestion levels. As a result, this process identified the periods when the port is most congested and the centroids of these clusters can be used to predict the behavior of port congestion levels. These indicators provide resources for better management and can motivate actions such as the redistribution of ship loading and unloading locations, improving the port performance measurement.

1. Introduction

Maritime trade represents around 90% of the global volume trade. Therefore, the port's performance is crucial to sustaining economic growth, (AbuAlhaol, Falcon, Abielmona, & Petriu, 2018). However, this increase in maritime trade has produced an impact on the Ports efficiency in some regions. In December 2019, for example, ships operated liquid chemical bulk in the Port of Santos had to wait more than 10 days for a docking opportunity, causing losses around US$ 35.000 per day for each ship, (Rosssi, 2019).

This situation is called port congestion, in which vessels must wait at areas close to the Ports for load or unload. In most cases, the port's capacity does not correspond to the demand, and the vessels, generally, must wait at anchorage areas before accessing the port, (MarineTraffic, 2020). This impact is not restricted to any part of the world, it also affects ports on Asia, North Africa, Northern Europe, and United States, (Saeed, Song, & Andersen, 2018).

Port congestion is an important issue from an economic and efficiency point of view. Because it results, not only, in longer waiting times and low service levels for vessels, but it also contributes to the decrease in competitiveness and demand (Saeed, Song, & Andersen, 2018). Understanding the aspects that influence congestion is essential for port management. However, traditional traffic analyzes are, generally, carried out through surveys that include: visual observations, radar, and aerial

Galoá { Este trabalho foi publicado utilizando o Galoá proceedings

photographs, being extremely costly, (Zhang, Men, & Fang Fwa, 2019).

Currently, there are advanced methods for collecting vessel traffic data, such as Vessel Traffic Services (VTS) and Automatic Identification System (AIS). The VTS is a maritime traffic monitoring system established and operated by port authorities. This technology uses radar, closed-circuit television, and high-frequency radiotelephony to track ship movements in a limited geographical area. However, the coverage of the VTS system is quite limited, as well as being expensive to access and process the data, since they are confidential and protected by maritime authorities. Also, the data information is limited and excludes important characteristics of the vessels, such as the type of vessel and draft (Meng, Weng, & Li, 2014).

The second type of data is the Automatic Identification System (AIS). AIS is a vessel tracking system that provides regular ship data updates. Static information (name of the vessel, type of vessel, length, breadth, etc.) and dynamic information (speed, navigation situation, direction, position, etc.) of the vessels can be exchanged electronically between AIS receiving stations (onboard, on land, or by satellite), (AbuAlhaol, Falcon, Abielmona, & Petriu, 2018). AIS data does not have the limitations of VTS data and, due to its informative integrity, can be used to analyze incidents, such as ship collisions. This data ensures greater reliability of navigation information and comes the analysis of maritime traffic more accurate, (Zhang, Men, & Fang Fwa, 2019), and (Meng, Weng, & Li, 2014).

However, most studies performed with AIS data have focused on specific areas such as monitoring, tracking, and security of ships, accident prevention, including collision risks, noise levels, or ship emissions, (Shelmerdine, 2015). Few researches employing AIS data have explored marine traffic congestion. Inspired by the work of Craighead et al. (2007), AbuAlhaol et al. (2018) proposed three "Big Data-Driven" indicators to measure the marine traffic congestion: i) the spatial density of seaports; ii) the spatial complexity; and iii) the average waiting time for ships, (Craighead , Blackhurst, Rungtusanatham , & Handfield, 2007), and (AbuAlhaol, Falcon, Abielmona, & Petriu, 2018).

From these indicators, the k-means clustering technique was used to identify the months of the year in which the selected ports were more or less congested. These Port Congestion Indicators (PCIs) provide resources for management and can motivate actions such as the redistribution of ship loading and unloading locations and better anchorage planning, (AbuAlhaol, Falcon, Abielmona, & Petriu, 2018).

This work employs a variation of the methodology proposed by AbuAlhaol et al. (2018) to analyze the port congestion level of the Port of Rio de Janeiro. Over one year, AIS data were collected and analyzed to calculate Port Congestion Indicators (PCIs): i) Spatial Concentration; ii) Spatial Density; iii) Average Service Time. Machine learning techniques were applied to these indicators to identify the weeks in which the port is most congested, allowing predict the future status.

2. Database

The Safety of Life at Sea (SOLAS) Convention published in 2002 by the International Maritime Organization (IMO) required that: i) all marine vessels over 300 GT (gross tonnage) on an international voyage; ii) all cargo vessels greater than 500 GT; iii) all passenger vessels irrespective of size, to be fitted with an Automatic Identification System (AIS) as standard by 2004, (Shelmerdine, 2015).

The Automatic Identification System (AIS) is a vessel tracking system that provides regular updates on a vessel's movement and other relevant ship voyage data to other parties, (AbuAlhaol, Falcon, Abielmona, & Petriu, 2018). The AIS was developed to avoid ship collision accidents, and it has been used in maritime transportation for over two decades. Originally, AIS data was highly regionalized and difficult to collect, because AIS communication was limited to very-high-frequency (VHF) radio wave range, which only covered 10–20 nautical miles, (Yang , Wang, & Jia, 2019).

Since 2008, satellites equipped with AIS receivers have been able to receive AIS data transmitted by onboard AIS transceivers worldwide. At present, AIS data can be easily collected from commercial websites that provide access to AIS databases. With the constantly improving quality and completeness of AIS data, the applications of AIS data have expanded from navigation safety to include many other aspects too, (Yang , Wang, & Jia, 2019).

The AIS system uses a transponder that transmits and receives in VHF. It also includes a GPS receiver that records the position and movement details. Transmission and reception are carried out continuously and autonomously, and the use of this

Galoá { Este trabalho foi publicado utilizando o Galoá proceedings

technology does not entail removing the preexisting systems, (Serry, 2018).

The information contained in each AIS data can be divided into the following two main categories: i) Dynamic Information (such information is automatically transmitted every 2 to 10 seconds depending on the vessel's speed and course while underway and every 6 minutes while anchored from vessels equipped with Class A transponders.); ii) Static & Voyage related Information (such information is provided by the subject vessel's crew and is transmitted every 6 minutes regardless of the vessel's movement status.), (MarineTraffic, 2020).

The database used in this work is the AIS messages transmitted by ships on the region of the port of Rio de Janeiro. The data consists of the period of January 2018 to April 2018 and September 2018 to March 2019. To become the analysis more accurate, the weeks in which AIS data were not transmitted, were removed.

3. Methodology

This section describes the methodology used in this work. The first step is to calculate the geospatial algorithms: convex hull area, geolocation area, and average vessel proximity. Then, these algorithms will be used to calculate the Port Congestion Indicators (PCIs): spatial concentration, spatial density, and average service time. Finally, machine learning techniques were employed to cluster these indicators, to extract useful information.

3.1. Geospatial Algorithms

A. Convex Hull Area

The convex hull of a set of points is the smallest convex set that contains the points, it is a fundamental construction for mathematics and computational geometry, (Barber, Dobkin, & Huhdanpaa, 1996). Quickhull is an algorithm for computing convex hulls that takes a divide-and-conquer approach. The idea is to partition the problem into subproblems of roughly equal size, solve each subproblem recursively, and finally combine the individual results into a whole solution, (Mucke, 2009).

The convex hull of a set of finite planar points is a polygon. Since any point in the convex hull can be expressed as a convex combination of its vertices, for simplicity, we use the phrase “convex hull” to mean “the set of vertices of the convex hull”, (Nguyen & Le, 2015). The port convex hull area is defined as the area that encloses all vessels in the smallest perimeter fence, (AbuAlhaol, Falcon, Abielmona, & Petriu, 2018).

The input to the convex hull algorithm is the latitude and longitude of the vessels from AIS static and dynamic messages. The first step is to filter: a) the latitude and longitude of the port of Rio de Janeiro, b) the navigational status moored or in anchor, and c) the speed over ground below 5 knots. The navigational status is manually set by the crew and there are sixteen statuses (e.g., 1 is for in anchor and 5 is for moored). The speed over ground is in nautical miles per hour (knots), and Abualhaol et al. (2018) considered the messages with a reported speed below 5 knots.

The convex hull area and vertices were calculated using Python’s SciPy open-source library, and Figure 1 was generated on Map Polygon Tool. As a result, throughout January 2018 to April 2018 and September 2018 to March 2019, the convex hull area was $138{,}09\ Km^2$. The detailed process over the entire period is provided on Algorithm 1, see Table 1.

Table 1: Algorithm 1 about Convex Hull Area.

Algorithm 1: Convex Hull Area
Data: AIS static and dynamic messages. **Filter:** a) Latitude and longitude of the port of Rio de Janeiro; b) Navigational status moored or in anchor; c) Speed over ground below 5 knots. **Input:** Latitude and longitude of the vessels. **Initialization:** 1. Conversion of latitude and longitude to x and y. 2. Calculation of the convex area. **Output:** Convex Area in Km^2.

Figure 1: Convex hull area of the port of Rio de Janeiro.

Galoá { Este trabalho foi publicado utilizando o Galoá proceedings

B. Geolocation Area

The geolocation area is related to the geohash system invented in 2008 by Gustavo Niemeyer. Geohash is an address code that reduces two-dimensional latitude and longitude information to a unique one-dimensional string for each point on earth with a given precision. This system uses a string to represent longitude and latitude coordinates and does not represent a point, but a region, (Xiang, 2019).

It is a geocoding method to encode geographic coordinates into a short string of digits and letters delineating an area on a map, which is called a cell, with varying resolutions. The more characters in the string, the more precise the location. Geohashes use Base-32 alphabet encoding, the first character in a geohash identifies the initial location as one of the 32 cells. This cell will also contain 32 cells, and each one of these will contain 32 cells (and so on repeatedly). Adding characters to the geohash sub-divides a cell, effectively zooming in to a more detailed area, (PubNub).

The precision factor determines the size of the cell. For instance, a precision factor of one creates a cell 5,000km high and 5,000km wide, a precision factor of five creates a cell 4.89km high, and 4.89km wide, and a precision factor of seven creates a cell 153m high and 153m wide, (PubNub). Abualhaol et al. (2018) use precision-7 (PREC = 7) geohashes, to calculate the geohash area.

In our work, we split the convex area into cells of $153m$ high and $153m$ wide (0,023409 Km^2), and the coordinates of these cells were used to identify which area is been used by the ships. Each cell can only be activated once in the period if there is at least one vessel within the coordinates of that cell. Then, to calculate the geolocation area, we considered the area of each activated cell over the period. The detailed process of each week is provided on Algorithm 2, see Table 2

Table 2: Algorithm 2 about Geolocation Area.

Algorithm 2: Geolocation Area
Data: AIS static and dynamic messages.
Input: Latitude and longitude of the vessels.
Initialization:
1. Split the convex area into cells of $153m$ high and $153m$ wide (0,023409 Km^2).
2. Identify which cells is been used by the vessels over the period.
3. Calculate the area of the activated cells over the period.
Output: Geolocation area in Km^2.

C. Average Vessels Proximity

When working with latitude and longitude, it is sometimes helpful to calculate distances between points. But simple Euclidean distance doesn't cut it since we have to deal with a sphere, or an oblate spheroid to be exact. So, we have to take a look at geodesic distances, (Janakiev, 2018). The geodesics on a surface are, locally at least, the curves of the shortest distance on the surface between any two points on that surface. In geodesy, the geodesic is well understood to refer to the shortest surface distance between two points on the surface of the ellipsoid, or synonymously the spheroid, (Thomas & Featherstone, 2005). There are various ways to handle this calculation as The Great-circle distance, the Haversine formula, and Vincenty's formulae.

The great circle or orthodromic distance is the shortest distance between any two points on a sphere measured along a path on its surface. Because the geometry of the sphere is different from ordinary Euclidean geometry, the equations for distance take on a different form. The distance between two points in Euclidean space is the length of a straight line from one point to the other. On a sphere, however, there are no straight lines. In non-Euclidean geometry, straight lines are replaced by geodesics. On the sphere, geodesics are the great circles, (Porcu, Bevilacqua, & Genton, 2016).

The haversine formula is a method used to calculate the distance from one place to the destination. This formula calculates the distance between two points (longitude and latitude) based on the length of the straight line. The haversine formula is commonly used in navigation problems because it can provide a large circle distance between two points on the surface of the globe regardless of the height of the hill and the depth of the valley on the surface of the earth, (Rezania Agramanisti & Febriyanti, 2020).

Vincenty's formulas are typically used for the calculation of the direct and inverse geodetic problems. The direct problem is: given the geodetic latitude and longitude (φ_1, λ_1) of a point on the surface of the ellipsoid, along with the starting azimuth α_1 and geodesic distance s, find the finishing point (φ_2, λ_2) and (reverse) azimuth α_2 of the geodesic at (φ_2, λ_2). The inverse problem is: given two points (φ_1, λ_1) and (φ_2, λ_2) on the surface of the ellipsoid, find the geodesic distance s between them, and the forward azimuth α_1 and reverse azimuth α_2 of the geodesic at (φ_1, λ_1) and (φ_2, λ_2) respectively, (Thomas & Featherstone, 2005)

Galoá { Este trabalho foi publicado utilizando o Galoá proceedings

The average vessel proximity was calculated using Phyton's Pyproj open-source library, which performs forward and inverses geodetic algorithms. The input to the algorithm is the latitude and longitude of the vessels. The first step is a combination, without repetition, of the vessel's coordinates. Then, the algorithm calculates the distance between the coordinate pairs. The interquartile range (IQR) method, was applied, on the results, to remove the outliers. The detailed process of each week is provided on Algorithm 3, see Table 3.

Table 3: Algorithm 3 about Average Vessel Proximity

Algorithm 3: Average Vessel Proximity
Data: AIS static and dynamic messages. **Input:** Latitude and longitude of the vessels. **Initialization:** 1. Combination, without repetition, of the vessel's coordinates. 2. Calculation of the distance between the coordinate pairs. 3. Interquartile range (IQR) method to remove outliers. 4. Calculation of the average vessel's proximity. **Output:** Average vessel proximity in Km

3.2. Port Congestion Indicators

A. Spatial Concentration

Spatial Concentration (SC) is the normalized port congestion indicator, to measure the spatial distribution of ships within the convex area. It is calculated by dividing the Convex Hull Area (Convex Area) by the Average Vessels Proximity (Δ), as presented in Algorithm 1 and Algorithm 3, respectively.

Equation 1 presents the analytical formulation of the Spatial Concentration Indicator where the subscript i indicates the aggregation period from a set of all periods (in this work i will be a week index and I will be a set of 11 months in 2018 and 2019, $i \in I$.). The Spatial Concentration is normalized on the maximum historical value (i.e., $max\ \{SC_i\}$, for $i \in I$) and therefore is unitless in the range $[0, 1]$.

We considered that spatial concentration increase as the relationship between the convex hull area and average vessel proximity grows. For the same average proximity but with a larger convex area the spatial concentration increase.

Equation 1: Spatial Concentration Indicator

$$SC_{(i)} = \frac{Convex\ Area_{(i)}/\Delta_{(i)}}{\max\limits_{i \in I} \{Convex\ Area_{(i)}/\Delta_{(i)}\}}$$

B. Spatial Density

Spatial Density (SD) is the normalized port congestion indicator, to measure the area that is been used by the vessels. It is calculated by dividing the Geolocation Area (Geo Area) by the Convex Hull Area (Convex Area), as presented in Algorithm 2 and Algorithm 1, respectively.

We considered that spatial density increase as the relationship between the geolocation area and the convex area grows. For the same convex area but with a larger geolocation area the spatial density increase.

Equation 2 presents the analytical formulation of the Spatial Density Indicator. It is normalized by the maximum reachable value in the periods of interests (i.e., I), and therefore is unitless in the range $[0;\ 1]$.

Equation 2: Spatial Density Indicator

$$SD_{(i)} = \frac{Geo\ Area_{(i)}/Convex\ Area_{(i)}}{\max\limits_{i \in I} \{Geo\ Area_{(i)}/Convex\ Area_{(i)}\}}$$

C. Average Service Time

The third Port Congestion Indicator (i.e., Average Service Time) represents the average time needed by vessels to enter, load/offload, and exit the port over the period. We define t_n as the time needed for a vessel to get served and leave the port, N_i the number of unique vessels (based on the reported MMSI numbers) in the i_{th} aggregation period.

The Average Service Time (AST) is the normalized average time for all cargo vessels in the i_{th} period as given in Equation 3. It is normalized on the maximum value and therefore is unitless in the range $[0;\ 1]$.

Equation 3: Average Service Time Indicator

$$AST_{(i)} = \frac{\frac{1}{N_i}\ x\ \sum_{n=1}^{N_i} t_n}{\max\limits_{i \in I} \left\{\frac{1}{N_i}\ x\ \sum_{n=1}^{N_i} t_n\right\}}$$

Galoá { Este trabalho foi publicado utilizando o Galoá proceedings

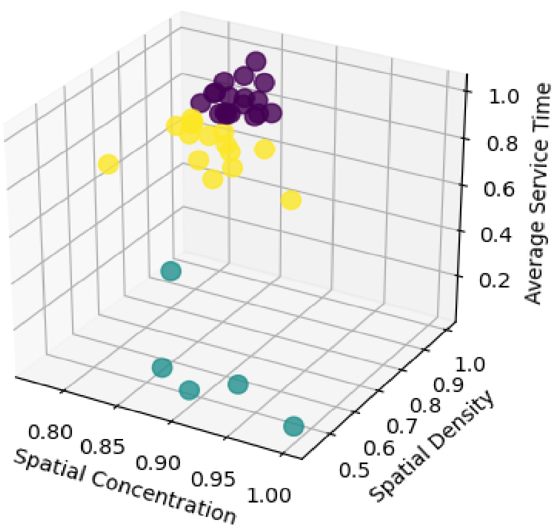
3.3. Knowledge Discovery in Database

Knowledge discovery in database (KDD) is the nontrivial extraction of implicit, previously unknown, and potentially useful information from data. The approaches are taken, however, are quite diverse. Most are based on machine learning methods that have been enhanced to better deal with issues particular to discovery in databases (Frawley, Piatetsky-Shapiro, & Matheus, 1992). The grand challenge of KDD is to automatically process large quantities of raw data, identify the most significant patterns, and present these as knowledge appropriate for user's goals (Matheus, Chan, & Piatetsky-Shapiro, 1993).

To extract useful information, one of the most popular unsupervised machine learning algorithms is the k-means, for clustering. The k-means procedure consists of simply starting with k groups each of which consists of a single random point, and thereafter adding each new point to the group whose mean the new point is nearest. After a point is added to a group, the mean of that group is adjusted to take account of the new point. Thus, at each stage, the k-means are, in fact, the means of the groups they represent (MacQueen, 1967).

In our work, the k-means is used to cluster the port congestion indicators: spatial concentration, spatial density, and average service time. We used the Scikit-Learn open-source library to perform the k-means clustering with $K = 3$ to cluster the aggregated periods (weeks) into three clusters. The purple color cluster (HIGH) is the closest to the (1; 1; 1) which represents the highest congestion cluster (i.e., the closest to maximum PCI values). The yellow cluster (MEDIUM) has moderate level congestion and finally, the low congestion level cluster is identified in blue (LOW), see Figure 2.

Figure 2: Port Congestion Indicators Clustering (K = 3)



4. Results

In this section, we present the port congestion indicators results for the port of Rio de Janeiro. Table 4 represents the main results of the indicators over the period. The months were divided into four weeks, being 1 for the first week and 4 for the last. The weeks in which AIS data were not transmitted, were removed.

Table 4: Port Congestion Indicators of Rio de Janeiro.

Week	MM/YY	SC	SD	AST	k-means
3	Jan/2018	0,87	0,91	0,85	HIGH
4	Jan/2018	0,80	0,83	0,77	MEDIUM
1	Feb/2018	0,87	0,88	0,71	MEDIUM
4	Feb/2018	0,84	0,63	0,36	LOW
1	Mar/2018	0,87	0,47	0,11	LOW
2	Mar/2018	0,84	0,94	0,95	HIGH
3	Mar/2018	0,84	0,85	0,78	MEDIUM
1	Apr/2018	0,83	0,90	0,89	HIGH
3	Apr/2018	0,77	0,69	0,69	MEDIUM
1	Sept/2018	0,82	0,84	0,89	HIGH
3	Sept/2018	0,80	0,88	0,77	MEDIUM
3	Oct/2018	0,82	0,92	0,87	HIGH
4	Oct/2018	0,82	0,87	0,73	MEDIUM
1	Nov/2018	0,83	0,88	0,69	MEDIUM
3	Nov/2018	0,80	0,99	0,80	HIGH
1	Dec/2018	0,84	0,93	0,84	HIGH
3	Dec/2018	0,94	0,46	0,14	LOW
2	Jan/2019	0,90	0,46	0,06	LOW
3	Jan/2019	0,82	0,94	0,78	HIGH
1	Feb/2019	0,84	1,00	0,88	HIGH
4	Feb/2019	0,91	0,82	0,59	MEDIUM
2	Mar/2019	0,85	0,83	0,65	MEDIUM
3	Mar/2019	1,00	0,43	0,07	LOW

Both Figure 2 and Table 4 represents the port congestion indicators from January 2018 to April 2018 and September 2018 to March 2019. Most of the months have weeks which is identified as the most severe congested cluster (spatial concentration, spatial density, and average service time close to 1,1,1). Few weeks presents low congestion levels, where may have been some technical issues receiving AIS data in Feb/2018, Mar/2018, Dec/2018, Jan/2019, and Mar/2019.

Nevertheless, both indicators of spatial concentration and spatial density did not show significant variations, and such indicators could

Galoá { Este trabalho foi publicado utilizando o Galoá proceedings

motivate the port of Rio de Janeiro authority to take special consideration in the high congested weeks redistributing the anchored vessels and widen the port area to decrease those indicators. On the other hand, to decrease the average service time, it may need to allocate more resources to speed up loading/unloading ships.

In general, the port of Rio de Janeiro is highly congested, and to decrease those indicators and make it more competitive is important to consider a redesign of operations and internal processes as well as continuing to monitor port performance.

5. Conclusions

Three geospatial algorithms: geohash area, convex hull area, and vessel average proximity were used to calculate the Port Congestion Indicator (spatial concentration, spatial density, average service time) from January 2018 to April 2018 and September 2018 to March 2019.

The k-means clustering with $K = 3$ was utilized to characterize congestion levels of the port of Rio de Janeiro into low, medium, and high. As a result, around 47%, over the period, the port has high congestion (port congestion indicators close to 1,1,1) and a 39% medium. The centroids of these clusters could be, also, used as the basis to predict the behavior of the port in future weeks.

The k-means clusters and the port congestion indicators bring actionable information to the port of Rio de Janeiro to understand the aspects that influence port congestion. However, it is important to improve the proposed indicators for a reviewed advanced model.

The proposed Indicators are reactive and based on historical AIS data. Nonetheless, it can be applied provocatively to classify port congestion levels based on a real-time AIS data stream. Apache Spark may be required to fast and distributed engine for largescale data processing. It provides distributed machine learning capabilities and can be reconfigured to enable real-time data processing.

Might be interesting to compare different ports for port authorities and stakeholders to start investigating why port congestion is occurring. Is the port overutilized and should its operation be improved or are there other reasons such as weather conditions or internal processes that need to be better accounted for?

This congestion is increasing costs for shippers and importers on the port of Rio de Janeiro, and the model presented could help other ports in the region. Few studies, in Latin America, of port congestion, were developed, thereby it is a contribution that can be useful for neighboring ports since the methodology can be easily applied to other study areas.

These surveys expand knowledge in the field since the quantification of port congestion provides resources to port authorities and stakeholders for better management, improving the port logistics operations, and reducing the costs.

References

AbuAlhaol, I., Falcon, R., Abielmona, R., & Petriu, E. (2018). Mining Port Congestion Indicators from Big AIS Data. *2018 IEEE International Conference on Big Data (Big Data)*, 2580-2590. doi:10.1109/IJCNN.2018.8489187

Barber, C., Dobkin, D., & Huhdanpaa, H. (1996). The Quickhull Algorithm for Convex Hull. *ACM Trans. Math. Softw., 22*(4), 469–483. doi:10.1145/235815.235821

Craighead, C., Blackhurst, J., Rungtusanatham, M., & Handfield, R. (2007). The Severity of Supply Chain Disruptions: Design Characteristics and Mitigation Capabilities. *Decision Sciences Journal, 38*(1), 131-156. doi:10.1111/j.1540-5915.2007.00151.x

Frawley, W., Piatetsky-Shapiro, G., & Matheus, C. (1992). Knowledge Discovery in Databases: An Overview. *AI Magazine, 13*(3), 57-70.

Janakiev, N. (2018, March 09). *Calculate Distance Between GPS Points in Python.* Retrieved from Parametric Thoughts: https://janakiev.com/blog/gps-points-distance-python/

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Berkeley Symposium on Mathematical Statistics and Probability, 1*, 281-297. Retrieved from https://projecteuclid.org/euclid.bsmsp/1200512992

MarineTraffic. (2020, 01 20). *MarineTraffic Help.* Retrieved from Port Congestion in the Port Details Page: https://help.marinetraffic.com/hc/en-us/articles/115003716171-Port-Congestion-in-the-Port-Details-Page

Matheus, C., Chan, P., & Piatetsky-Shapiro, G. (1993). Systems for Knowledge Discovery in Databases. *IEEE Transactions On*

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings

Knowledge And Data Engineering, 5, 903-913.

Meng, Q., Weng, J., & Li, S. (2014, January). Analysis with Automatic Identification System Data of Vessel Traffic Characteristics in the Singapore Strait. *Transportation Research Record: Journal of the Transportation Research Board, 2426* (1), 33-43. doi:10.3141/2426-05

Mucke, E. (2009). Quickhull: Computing Convex Hulls Quickly. *Computing in Science & Engineering, 11*(5), 54-57. doi:10.1109/MCSE.2009.136

Nguyen, K., & Le, D. (February de 2015). A Convex Hull algorithm for solving a location problem. *RAIRO-Oper. Res., 48*(3), 589-600. doi:10.1051/ro/2014058

Porcu, E., Bevilacqua, M., & Genton, M. (2016). Spatio-Temporal Covariance and Cross-Covariance Functions of the Great Circle Distance on a Sphere. *Journal of the American Statistical Association, 111*(514), 888-898. doi:10.1080/01621459.2015.1072541

PubNub. (n.d.). *PubNub.* Retrieved 07 22, 2020, from What is Geohashing?: https://www.pubnub.com/learn/glossary/what-is-geohashing/

Rezania Agramanisti, A., & Febriyanti, D. (2020). Use of Haversine Formula in Finding Distance Between Temporary Shelter and Waste End Processing Sites. *Journal of Physics: Conference Series, 1500*(012104), 1-6. doi:10.1088/1742-6596/1500/1/012104

Rosssi, M. (16 de 12 de 2019). Navios esperam 10 dias para atracar em Santos e prejuízo ultrapassa US$ 10 milhões. *G1 Santos*. Fonte: https://g1.globo.com/sp/santos-regiao/porto-mar/noticia/2019/12/16/navios-esperam-10-dias-para-atracar-em-santos-e-prejuizo-ultrapassa-us-10-milhoes.ghtml

Saeed, N., Song, D.-W., & Andersen, O. (2018). Governance model for port congestion mitigation: A transaction cost perspective. *NETNOMICS: Economic Research and Electronic Networking, 19*, 159–178. doi:10.1007/s11066-018-9123-4

Serry, A. (2018). The automatic identification system (AIS): a data source for studying maritime traffic. *7th INTERNATIONAL CONFERENCE ON MARITIME TRANSPORT. Technological, Innovation and Research. MARITIME TRANSPORT ‘16*, 346-360. Retrieved from https://hal.archives-ouvertes.fr/hal-01724104/document

Shelmerdine, R. (2015, April). Teasing out the detail: How our understanding of marine AIS data can better inform industries, developments, and planning. *Marine Policy, 54*, 17-25. doi:10.1016/j.marpol.2014.12.010

Thomas, C., & Featherstone, W. (2005). Validation of Vincenty’s Formulas for the Geodesic Using a New Fourth-Order Extension of Kivioja’s Formula. *Journal of Surveying Engineering, 131*(1). doi:10.1061/(ASCE)0733-9453(2005)131:1(20)

Xiang, W. (2019). An Efficient Location Privacy Preserving Model based on Geohash. *6th International Conference on Behavioral, Economic and Socio-Cultural* (pp. 1-5). Beijing, China: IEEE. doi:10.1109/BESC48373.2019.8963346

Yang, D., Wang, S., & Jia, H. (2019). How big data enriches maritime research – a critical review of Automatic Identification System (AIS) data applications. *Journal Transport Reviews, 39*(6), 755-773. doi:10.1080/01441647.2019.1649315

Zhang, L., Meng, Q., & Fang Fwa, T. (2019, September). Big AIS data-based spatial-temporal analyses of ship traffic in Singapore port waters. *Transportation Research Part E: Logistics and Transportation Review, 129*, 287-304. doi:10.1016/j.tre.2017.07.011

Galoá { Este trabalho foi publicado utilizando o **Galoá** proceedings