An Information Extraction of LAPAN Satellite's AIS Database for Ship Classifications on the Territorial Waters of Indonesia

Muazam Nugroho[1], Nurrochman Ferdiansyah[1], Dicka Ariptian Rahayu[1], Rizki Permala[1], Patria Rachman Hakim[1], Wahyudi Hasbi[1]

[1]Satellite Technology Center, National Institute of Aeronautics and Space (LAPAN), Indonesia

e-mail: muazam.nugroho@lapan.go.id

Received: 23-03-2021. Accepted: 06-05-2021. Publishied: 30-06-2021

Abstract

LAPAN participates in maritime surveillance with Automatic Identification System (AIS) payloads on the LAPAN-A2 and LAPAN-A3 satellites. AIS data consists of dynamic messages and static messages. A complete data set is needed to facilitate monitoring and analysis of ship traffic. This complete data contains the ship's identity (static message) and its navigation status (dynamic message). This data set can be obtained by processing data in the database. This study focuses on extracting information from the LAPAN satellite AIS database with combining AIS data from dynamic and static messages to get the number of vessels grouped by vessel type in the territory of Indonesia with coordinates 95BT / 141BT, -11LS / 6LU with the data period January to December 2019 monthly. The information extraction on AIS satellite data for the classification of vessel types has been completed. The result of the information extraction is the greater the period of static data, the lesser the ship type is unknown / null, with average percentage of null vessel type is 73.76%, 56.93%, and 39.58% for static data for a period of 1 month, 3 months, and 1 year respectively. The results also show cargo ships have the highest number of vessels, with the number of vessels shows a downward trend along the year of 2019.

Keywords: *AIS, database, LAPAN satellite, marine traffic*

1. Introduction

Indonesia is an archipelago country, with sea area larger than land area, making Indonesia one of the largest maritime countries in the world. The Indonesian sea area is one of the busiest shipping lanes in the world. Many ships passing through Indonesian territory, such as cargo ships, passenger ships, and fishing vessels. Large areas of the sea and dense maritime traffic require a good monitoring system to maintain security and reduce the level of problems or law violations at sea. The existence of a monitoring system on the movement of ships in Indonesia can reduce the level of the number of problems that occur in Indonesian seas (Judianto & Wahyudiono, 2014).

Automatic Identification System (AIS) is a technology used for monitoring of ship traffic around the world. AIS operates on the VHF channel with a frequency of 161,975 MHz and 162,025 MHz for the purpose of sending and receiving ship-specific information (ITU, 2012). Some of the information obtained from the AIS system includes Maritime Mobile Service Identity (MMSI), Position, Course, Call Sign, and AIS base station (Hu et al., 2012). AIS devices operate using a time division multiple access (TDMA) system, which makes it possible to send around 2000 reports per minute (Stupak, 2014). Messages sent by AIS devices from ships can be received by fellow ships, terrestrial-based AIS receiving stations, and satellite-based AIS receiving stations (space-borne AIS). The space-borne AIS has an advantage in the radius of receiving AIS messages because of its height (Ball, 2013). The National Aeronautics and Space Agency (LAPAN) participated in Indonesia's maritime surveillance by launching two micro satellites containing the AIS receiver, namely the LAPAN-A2 satellite in 2015 and LAPAN-A3 in 2016. The LAPAN-A2 satellite has a near-equatorial orbit with an altitude of 650 km while the LAPAN-A3 satellite has a sun-synchronous polar orbit with an altitude of 515 km. LAPAN-A2 with its orbit allows it to pass through Indonesian territory 14 times a day, while the polar orbit on the LAPAN-A3 satellite allows for global coverage (Karim et al., 2018). With these different types of orbit, the two satellites could complement each other's AIS data recorded by each satellite.

Several studies have been conducted regarding the use of AIS data. (Yitao et al., 2017) proposed a route extraction method based on satellite-AIS which includes data preprocessing, structural similarity calculation, clustering and route extraction. (Zhou et.al, 2019) presents a methodology for clustering ship behavior in an area and classifying ships into these clusters based on the static ship characteristics. (Kundakci & Nas, 2018) developed a desktop application to enable visualization of vessel movements to map the ship movements according to vessel types and enables analysts to evaluate waterway traffic density in the target area. The AIS system is used in numerous applications that can contribute to a significant increase in the safety of passengers, crew, cargo and ships, as well as the marine environment (Stupak, 2014).

This study focuses on extracting information from the AIS LAPAN database to obtain information on the number of vessels based on unique MMSI which are grouped according to the type of ship entering Indonesian waters with coordinates 95BT / 141BT, -11LS / 6LU in the period January to December 2019. The method used in this study focuses mainly on AIS data processing in the database, such as database merging, data cleaning, view creation, and data retrieval using SQL queries. The purpose of this paper is to provide an overview of the classification of ships passing in Indonesian territorial waters during a time period.

2. Methodology

The method used in this study focuses mainly on AIS data processing in the database, such as database merging, data cleaning, view creation, and data retrieval using SQL queries. The process of extracting information from the LAPAN satellite AIS database is described in the following flowchart:

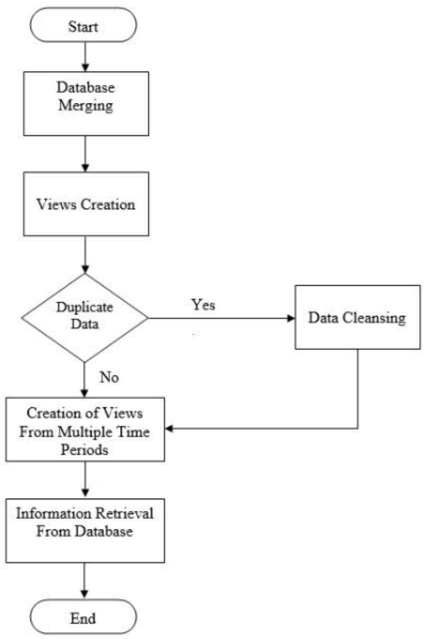

Figure 2-1: process of extracting information from the LAPAN satellite AIS database

It is explained in the IMO document (2002) that AIS messages transmitted by ships has four different types with different transmission times:

1. Static Information, which is information entered during the installation of AIS onboard and can also be changed if there is a change in the legal documents of the ship, such as a change in the name and ownership of the ship.
2. Dynamic Information, which is information that contains the ship's navigation status which will automatically change according to the data obtained from sensors connected to the AIS tool.
3. Sailing information, which is information entered and updated manually by the crew.
4. Short safety-related messages.

Details of the AIS information can be seen in the following figure:

Static information	Dynamic information	Voyage related information
Every 6 minutes and on request by a competent authority	*Dependent on speed and course alteration*	*Every 6 minutes, when data is amended or on request*
MMSI (Maritime Mobile Service Identity)	**Ship's position**	**Ship's draught**
Call sign and **name**	**Position Time stamp** in UTC	**Hazardous** cargo (type)
IMO Number	**Course over ground** (COG)	**Destination** ETA (Estimated Time of Arrival)
Length and beam	**Speed over ground** (SOG)	
Type of ship	**Navigational status** (underway, at anchor, moored...)	
Location of position fixing antenna	Rate of turn (ROT)	

Figure 2-2: information transmitted by ship (IMO, 2002)

In static message there is a field of MMSI (Maritime Mobile Service Identity) which is a unique code for the ship. The making of the MMSI code is regulated in the ITU-R M.585-8 document, it explains that the MMSI code has a nine-digit length with the format $M_1I_2D_3X_4X_5X_6X_7X_8X_9$ where the initial three digits represent the code of origin of the ship's country. Static message and shipping message are transmitted every 6 minutes or when there is a request from the authorities while for dynamic information the transmission interval is carried out according to the status and speed of the ship:

Ship's dynamic conditions	Nominal reporting interval
Ship at anchor or moored and not moving faster than 3 knots	3 min(1)
Ship at anchor or moored and moving faster than 3 knots	10 s(1)
Ship 0-14 knots	10 s(1)
Ship 0-14 knots and changing course	3 1/3 s(1)
Ship 14-23 knots	6 s(1)
Ship 14-23 knots and changing course	2 s
Ship >23 knots	2 s
Ship >23 knots and changing course	2 s

Figure 2-3: transmission data interval

2.1. Database Processing

As previously mentioned, LAPAN-A2 and LAPAN-A3 satellites carry the AIS receiver payload into orbit. LAPAN-A2 satellite in near-equatorial orbit receives AIS messages around the equatorial line, while with its polar orbit, LAPAN-A3 satellite receives AIS messages which are more evenly distributed around the world. The data stored in the satellite is routinely acquired by the earth station and stored in a database management system. The AIS data from these two satellites are stored in different databases. For the purposes of utilization, AIS data from separate databases are combined to complement each other in order to obtain more complete information.

The reason for using a database system to store data is that the database offers speed, accuracy, reporting, and thoroughness as its advantages (Suehring, 2002). To store satellite AIS data into a database, a Database Management System (DBMS) is required. MySQL is a DBMS with a relational system, also known as a relational database management system (RDBMS). Relational databases store data in separate tables (MySQL, 2021).

In the database, AIS messages are separated into tables based on the type of information. Dynamic messages (which have position coordinate fields) are entered in the Position Report table, whereas static messages are entered in the static table. The purpose of this paper is to obtain information on ship types in a certain area and at a certain time, then the information needed is the position data, the time contained in the position report table and the ship type data in the static table.

The method used in this study focuses on processing the AIS database for the LAPAN-A2 and LAPAN-A3 satellites. A summary of the research methods is described below:

2.1.1. Database Merging

The AIS database for the LAPAN-A2 and LAPAN-A3 satellites is combined into a new database. The table format in the database is made exactly the same as the database before it is merged. What is different is that the new database contains AIS messages originating from the LAPAN-A2 and LAPAN A3 satellites, and the information period is grouped into monthly period.

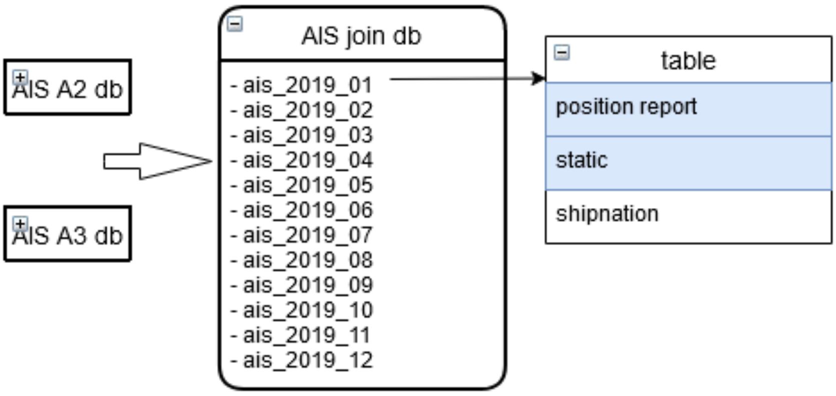

Figure 2-4: Merging of the AIS database for the LAPAN-A2 and LAPAN-A3 satellites

2.1.2. Views Creation

In a MySQL database, a view can be defined as a virtual table. The contents of this table can come from other tables or a combination of several tables. The purpose of creating a View is to make it easier to write queries and speed up data display, especially if the query is repeated (DuBois, 2014). In this study, making a view from each monthly database aims to retrieve the ship type data in the static table into the location message row in the Position Report table.

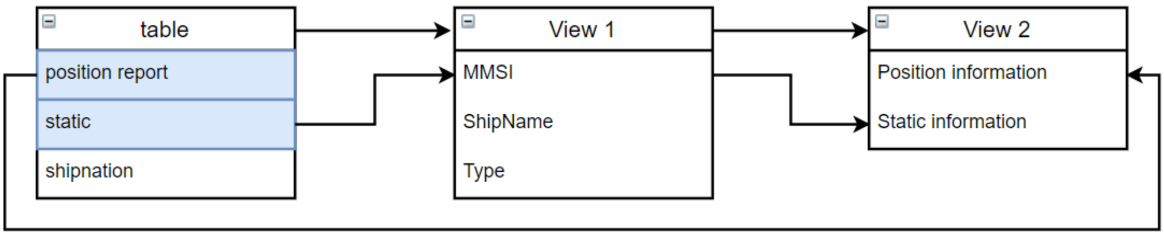

Figure 2-5: Creation of Views on the LAPAN satellite AIS database

There are 2 views created. The first view (View 1) contains a list of ships with unique data on MMSI, ship name, and vessel type. This data is obtained from a static table. The second view (View 2) contains a combined data from the Position Report table with data from View 1 based on MMSI. This combination will provide data on the ship's position complete with the name and type of ship. Data from View 2 is what data retrieval and analysis is carried out.

2.1.3. Data Cleansing

Filter queries use distinct on a static table to filter MMSI, ship name and type so there is no duplication, in other words, only 1 unique MMSI value represents 1 ship. Furthermore, the max (time) filter is applied in the Position Report table to obtain a more valid vessel type.

2.1.4. Creation of Views from multiple time periods

Static data in the database is usually incomplete, which is no static data from the MMSI in the Position Report data, which will cause the ship type to be unknown / NULL. This could be due to not receiving messages from the ship or errors during the transmission of AIS messages.

Table 2-1: Period of analyzed Position Report data and static data

no	Position Report	Static Data 1 month	Static Data 3 months	Static Data 1 year
1	Jan-19	Jan-19	Jan, Feb, Mar 2019	Jan - dec 2019
2	Feb-19	Feb-19		
3	Mar-19	Mar-19		
4	Apr-19	Apr-19	Apr, May, Jun 2019	
5	May-19	May-19		
6	Jun-19	Jun-19		
7	Jul-19	Jul-19	Jul, Aug, Sep 2019	
8	Aug-19	Aug-19		
9	Sep-19	Sep-19		
10	Oct-19	Oct-19	Oct, Nov, Dec 2019	
11	Nov-19	Nov-19		
12	Dec-19	Dec-19		

In this study, a view was made of creating views by taking static data from several different time periods, namely the period of one month, 1 quarter, and 1 year. The purpose of retrieving static data from several time periods is to see how much influence the static data period has on the completeness of the data. The initial hypothesis is that the greater the static data period, the more complete the information will be. Table 2-1 describes the data period analyzed.

To combine the data in separate tables, a field is needed in both tables. This field is the MMSI (Maritime Mobile Service Identity) field. Here are the relationships between tables to produce the desired information:

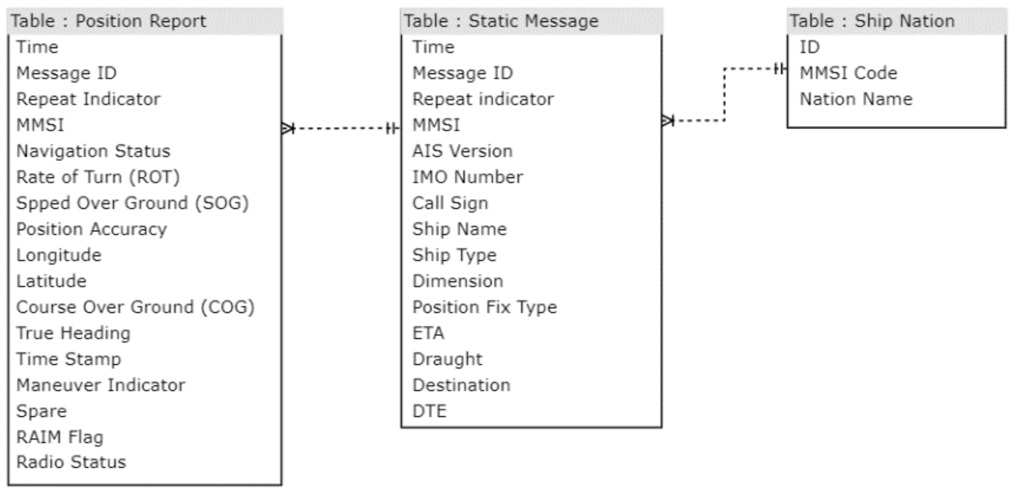

Figure 2-6: Relationships between tables in a database

2.2. Information Retrieval from the Database

Data retrieval is done by performing a MySQL query to calculate the number of unique MMSIs with time and region filters according to the needs. The following is the MySQL query used:

```
mysql>  SELECT  type, COUNT(*) FROM (select * from View2 where date(time) >='2019/01/01' and
        DATE(time)<='2019/01/31' and (lattd>=-11 and lattd<=6) and (longtd>=95 and
        longtd<=141) )
        AS A JOIN
        (select distinct mmsi, max(time) AS time from PositionReport where date(time)>=
        '2019/01/01' and DATE(time)<='2019/01/31' and (longtd>=95 and longtd<=141)
        and (lattd>=-11 and lattd<=6) group by mmsi)  AS B ON A.MMSI=B.mmsi AND
        A.Time=B.Time GROUP BY type;
```

This query will generate a number of unique MMSIs at a predetermined time and region with the vessel type of the MMSI.

3. Result And Discussion

After going through the data selection and classification process, the AIS satellite LAPAN-A2 and LAPAN-A3 data were obtained in the 95BT / 141BT, -11LS / 6LU areas in the period January to December 2019 which had the following data:

1. Ship Classification data with 1-month static data period

The following is the number of vessels by type in the period from January to December 2019 in the area 95BT / 141BT, -11LS / 6LU with static data for a period of 1 month:

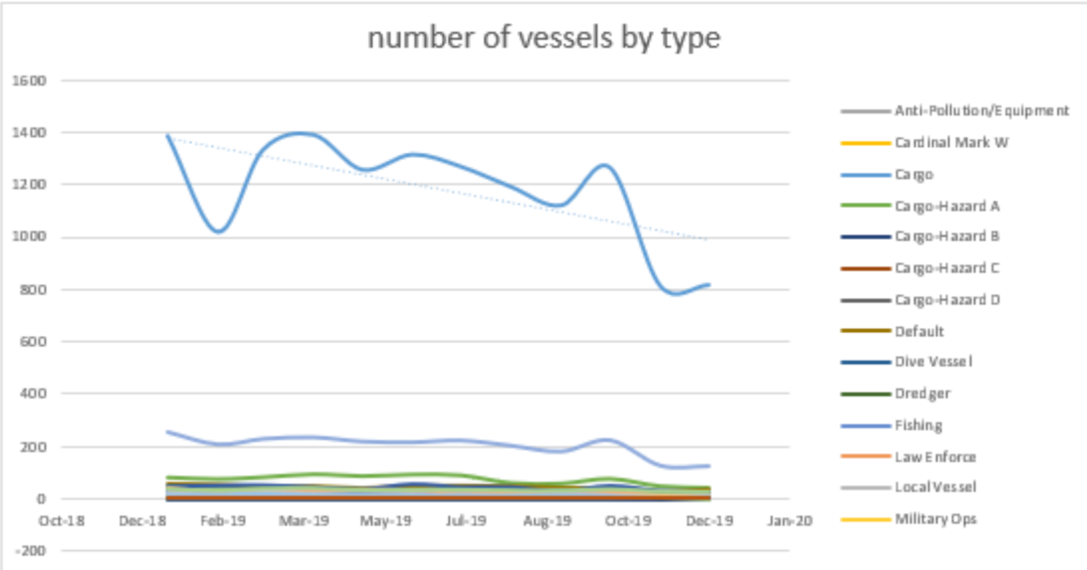

Figure 3-1: Graph of the number of vessels by type with a static data period of 1 month (without nulls)

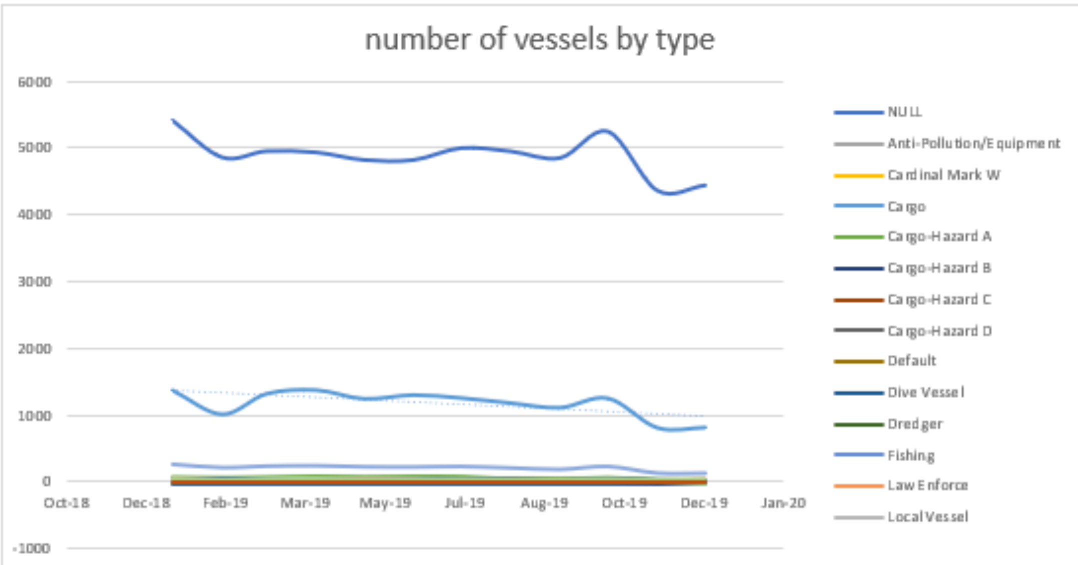

Figure 3-2: Graph of the number of vessels by type with a static data period of 1 month (with nulls)

From the graph above, it can be seen that the top 3 types of vessels are the unknown ship type (NULL), cargo type, and tanker type. Table 3-1 shows details of the number of unique MMSIs based on the 3 types of vessels:

Table 3-1: number of unique MMSIs from top 3 vessel type

		Ship Type	
month	**NULL**	**Cargo**	**Tanker**
Jan-19	5425	1386	251
Feb-19	4862	1021	204
Mar-19	4957	1334	225
Apr-19	4938	1390	231
May-19	4822	1257	215
Jun-19	4821	1315	212
Jul-19	5004	1269	219
Aug-19	4954	1192	199
Sep-19	4852	1122	177
Oct-19	5256	1265	220
Nov-19	4353	816	123
Dec-19	4434	820	121

2. Ship Classification data with 3-month static data period

The following is the number of vessels by type in the period from January to December 2019 in the area 95BT / 141BT, -11LS / 6LU with static data for a period of 3 month:

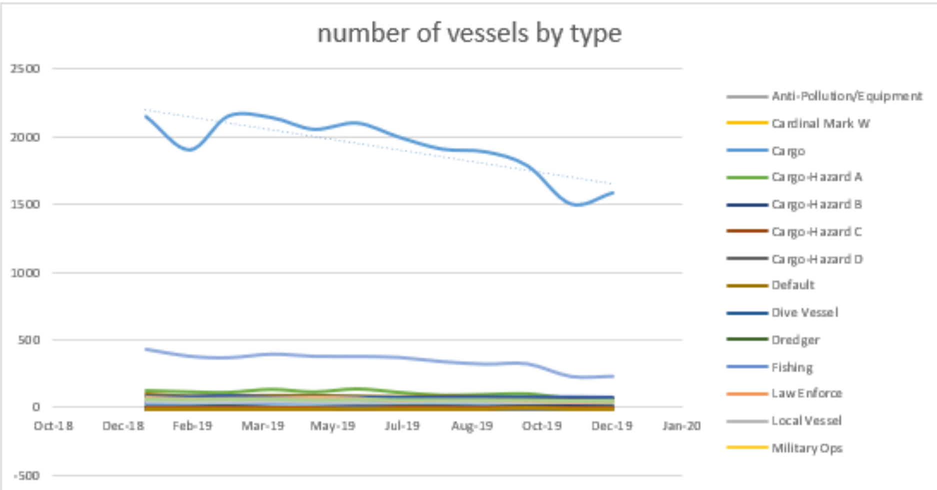

Figure 3-3: Graph of the number of vessels by type with a static data period of 3 month (without nulls)

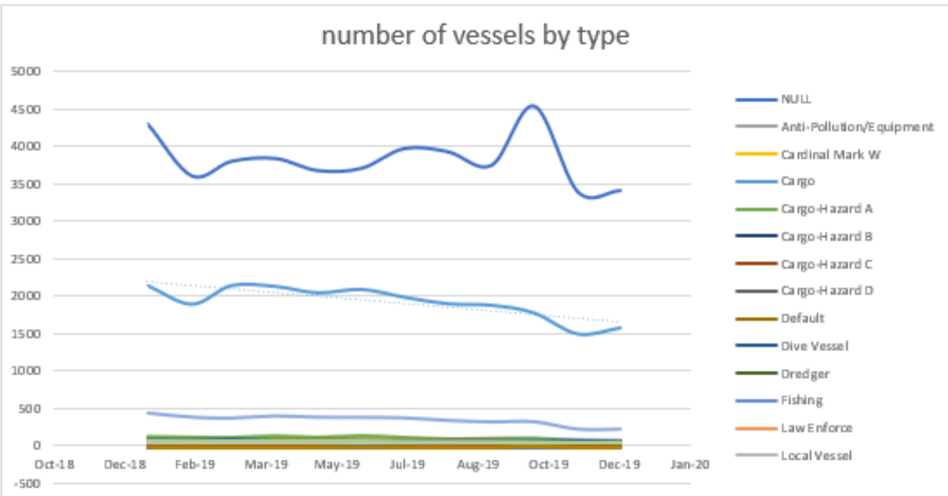

Figure 3-4: Graph of the number of vessels by type with a static data period of 3 month (with nulls)

From the graph above, it can be seen that the top 3 types of vessels are the unknown ship type (NULL), cargo type, and tanker type. Table 3-2 shows details of the number of unique MMSIs based on the 3 types of vessels:

Table 3-2: number of unique MMSIs from top 3 vessel type

		Ship Type	
month	**NULL**	**Cargo**	**Tanker**
Jan-19	4291	2148	428
Feb-19	3605	1902	378
Mar-19	3800	2152	367
Apr-19	3838	2140	393
May-19	3677	2054	377
Jun-19	3707	2100	376
Jul-19	3969	1996	369
Aug-19	3928	1908	338
Sep-19	3747	1888	320
Oct-19	4535	1785	322
Nov-19	3393	1504	230
Dec-19	3411	1586	231

3. Ship Classification data with 1-year static data period

The following is the number of vessels by type in the period from January to December 2019 in the area 95BT / 141BT, -11LS / 6LU with static data for a period of 1 year:

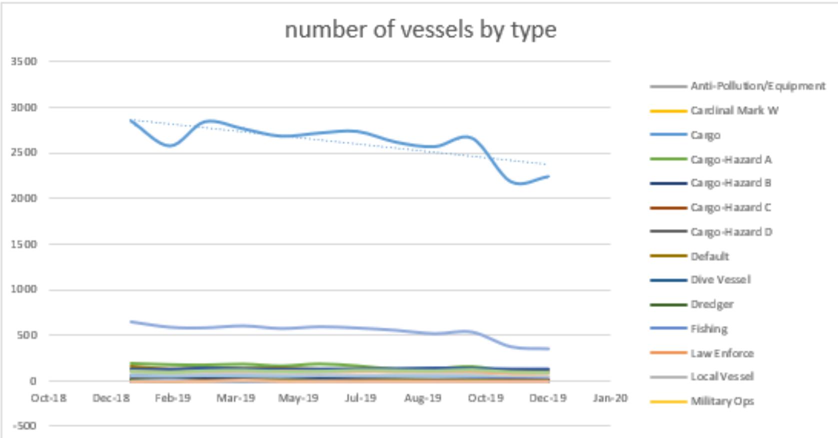

Figure 3-5: Graph of the number of vessels by type with a static data period of 1 year (without nulls)

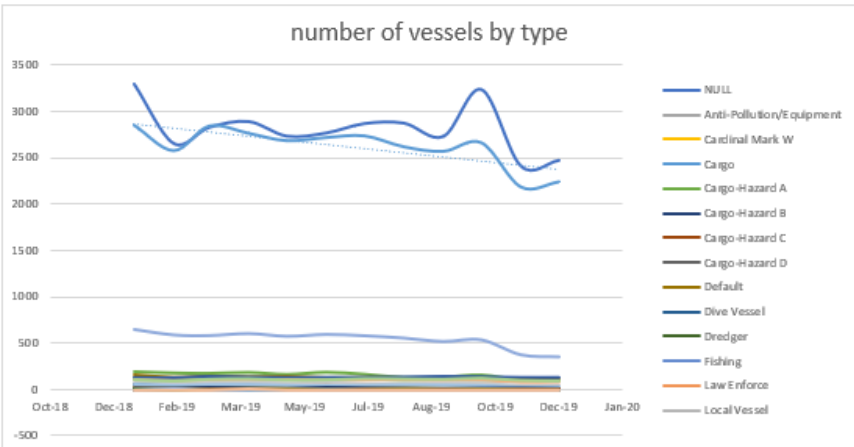

Figure 3-6: Graph of the number of vessels by type with a static data period of 1 year (with nulls)

From the graph above, it can be seen that the top 3 types of vessels are the unknown ship type (NULL), cargo type, and tanker type. Table 3-3 shows details of the number of unique MMSIs based on the 3 types of vessels:

Table 3-3: number of unique MMSIs from top 3 vessel type

	Ship Type		
month	**NULL**	**Cargo**	**Tanker**
Jan-19	3290	2845	651
Feb-19	2652	2574	591
Mar-19	2824	2837	584
Apr-19	2887	2759	606
May-19	2730	2680	576
Jun-19	2764	2714	596
Jul-19	2865	2732	582
Aug-19	2870	2615	556
Sep-19	2726	2566	518
Oct-19	3231	2658	536
Nov-19	2408	2185	371
Dec-19	2470	2240	348

4. Conclusion

By utilizing the AIS data in the database, we can get some statistical information on ships that are in Indonesian waters in the period from January to December 2019. Before this information can be obtained, it needs to be processed first in the database, one of which is by creating a view / relation table so that the data obtained to be complete.

In this study, information extraction on AIS satellite data for the classification of ship types has been completed. The result of extracting this information is that the greater the period of static data, the lesser the ship type is unknown / null, the average percentage of ship types null is 73.76% for a static data period of 1 month, 56.93% for a static data period of 3 months, and 39.58% for static data for a period of 1 year. In the period from January to December 2019, the largest type of ship was cargo, and it can be seen that the number of cargo ships has a downward trend along the year of 2019.

Acknowledgements

Authors sincerely appreciate Mr. Ir. Mujtahid, M.T. as the Head of Satellite Technology Center and LAPAN Satellite's operators for their support regarding AIS data used on this research.

Reference

A. Karim, R. Permala, M. Mukhayadi, W. Hasbi. LAPAN-A2 and LAPAN-A3 Automatic Identification System (AIS) Satellitte Data Correction using Interpolation and Extrapolation Method. Jurnal Teknologi Dirgantara vol. 16 no.2 December 2018. http://dx.doi.org/10.30536/j.jtd.2018.v16.a3049

Ball H., 2013. Satellite AIS for Dummies Special Edition. John Wiley & Sons Canada, Ltd. Mississauga. ISBN: 978-1-118-48494-4

DuBois Paul, 214. MySQL Cookbook, Third Edition. O'Reilly Media, Inc. Sebastopol, CA. ISBN: 978-1-449-37402-0

IMO: Recommendation on Performance Standards for a ship-borne Automatic Identification System (AIS). MSC 74(69) Annex 12. London: IMO, 2002

International Telecommunications Union. Interim Solutions for Improved Efficiency in the Use of the Band 156–174 Mhz by Stations in the Maritime Mobile Service ; International Telecommunications Union: Budapest, Hungary, 2012

Judianto, C. T., A. Wahyudiono. Automatic Identification System And Surveillance Technology For Indonesia Marine Security On Lapan-A2 Satellite. International Seminar of Aerospace Science and Technology II: 115 – 122, 2014

Kudakci B., Nas S. Mapping Marine Traffic Density By Using Ais Data: An Application In

The Northern Aegean Sea. Polish Maritime Research 4 (100) 2018 Vol. 25 page 49-58, 2018. https://doi.org/10.2478/pomr-2018-0131

MySQL. (2021). MySQL Documentation. Retreived March 19, 2021, from https://dev.mysql.com/doc/refman/8.0/en/what-is-mysql.html

Q. Hu, Yi Jiang, X. Sun, S. Zhang. Development of Automatic Indetification System Autonomous Positioning System. Journal Sensors, (2012). https://doi.org/10.3390/s151128574

Suehring Steve, 2002. MySQL Bible. Wiley Publishing, Inc. New York. ISBN: 0-7645-4932-4

T. Stupak. Influence of Automatic Identification System on Safety of Navigation at Sea. Journal of Marine Navigation and Safety of Sea Transportation Vol. 8 No. 3, 337-341, 2014. http://dx.doi.org/10.12716/1001.08.03.02

Yitao W., Lei Y., Xin S. Route Mining from Satellite-AIS Data Using Density-based Clustering Algorithm. 3rd International Symposium on Big Data and Applied Statistics, 2017. http://dx.doi.org/10.1088/1742-6596/1616/1/012017

Zhou Y, Daamen W, Vellinga T., Hoogendoorn S.P. Ship classification based on ship behavior clustering from AIS data. Ocean Engineering volume 175 1 March 2019 Pages 176-187, 2019. https://doi.org/10.1016/j.oceaneng.2019.02.005