Simulation of Computation Time of Ship Traffic Analysis using Data Clustering for a Vessel Traffic System

Rahimuddin[1], Mehdi Nakisa[2], Haryanti Rivai[1], Hasnawiyah Hasan[1], Husni Sitepu[1], Abdul Basir[3], Samsuddin[3], Albertha[4], Nenny[4]
[1]Hasanuddin University
[2]Islamic Azad University, Bushehr Iran.
[3]Politeknik Ilmu Pelayaran Makassar
[4]Politeknik Pelayaran Barombong, Indonesia
e-mail: rahimnav@unhas.ac.id

Abstract

Indonesian water territory is a crossroads of international shipping lanes connecting countries bordering the South China Sea in the north, countries in southern, and countries from western Indonesia. The shipping route goes through the Sunda Strait, Sumbawa Strait, Makassar Strait and Melaka Strait. The shipping lanes are grouped into three main routes called ALKI (Alur Laut Kepulauan Indonsia) I, II, and III. The increase in total number of ships per year with various capacities tends to increases the traffic in the lane and the risk of ship collision. Using AIS (Automatic Identification System) technology with a warning system required to prevent the effect of accidents. Through the Vessel Traffic System (VTS) receive, record, and analysis the traffic data in the covered area. Computation time become a critical issue for ship traffic assessment in the lane. A numerical simulation program developed and analysed the computation time of ship traffic analysis using cluster area method. The result showed a shorter computation time for a more number of cluster area. The computation time increase quadratically with an increase of ships number or objects.

Keywords: VTS, AIS, ship traffic, Computation Time, Data Clustering, DBSCAN

1. Introduction

Natural and human factors can be the main causes of marine accidents to occur. Natural factors such as bad weather that hit the ship beyond its capacity limit to overcome these disturbances, or ship accidents aground due to the very limited depth of the sea. Congested sea traffic can also affect a sea accidents when VTS receive inaccurate AIS data from the ships. Although humans cannot control nature, navigators can prevent accidents by using appropriate technology. However, accidents due to human error are the most common causes of accidents. The accident occurred due to the late in carrying out the navigation actions or taking inappropriate navigation actions.

Based on data from the Ministry of Transportation of Indonesia, 53,068 ships per year crossing the Sunda Strait, sailing from south to north in the South China Sea. The number of international shipping vessels crossing the Sunda Strait has increased while shipping from and to the port of Merak - Belawan has also increased. The domestic ship will have a crossing situation to the international shipping ships, resulted in a high risk traffic. The impact of the high congestion lane affects an increase of potential of sea

accidents. For this reason, Indonesian government proposed a separation of the international route of Sunda strait to Lombok strait to the IMO. The congestion of Lombok strait still lower with the number of ships passing is still around 36.773 ships per year, lower than the Sunda Strait. This separation can lower the congestion of Sunda strait and decrease the potential of sea accident.

With the use of AIS technology that has been regulated in IMO since in the early 2000s, monitoring of ship traffic in sea lanes in Indonesia is carried out by VTS under the marine transportation department. The ship's position is transmitted via the AIS system and the receiver system on the VTS can receive AIS data within the range of its receiver. VTS analyze and monitor the ship traffic to prevent congestion and sea accident. In monitoring ship traffic, VTS can take a warning based on ship collision prevention regulations in SOLAS, IMO in accordance with existing shipping lanes.

Various researches in the field of navigation safety, risk assessment of ship collision, and development of ship collision avoidance have been carried out by many researchers. The basic parameters to assess navigation safety were DCPA and TCPA. Along with the development of AIS technology and computing technology, the methods with extended parameters were developed, among others; ship length, relative speed, and relative ship direction to other ships are used in the development of ship collision prevention systems [1][2][3][4].

Computational capability in analyzing ship traffic is largely determined by the effectiveness of the algorithm used. In monitoring sea traffic, real-time data analysis must be carried out in a very limited time. The time needed to carry out a traffic analysis should be lower than the time required to take an immediate action by a navigator to prevent a marine accident. In 2017, Rong Zhen et. al using the DBSCAN algorithm in data clustering. This algorithm reduce the computation time required to analysis the traffic using the ship distance criteria with potential of collisions [5].

In 1997, Helmu Hilgert and Michael Baldauf developed a general model for risk assessment of ship collision situations with the addition of observation variables, such as radius of observation, ship speed, COLREG status, vessel length, weather visibility and sailing area [6]. Development of a risk assessment model using fuzzy logic and search algorithm by adding variables of ship speed and ship length [7]. The collision risk assessment model in the VTS coverage area uses the fuzzy logic method with four variables of ship speed, vessel length, wave conditions, and operating area [8].

The traffic analysis requires the accuracy of the data to produce the right analysis. The AIS data received by VTS from one ship is not syncron with the received data from others ships. The actual traffic conditions at the time of the traffic analysis conducted is distorted. One of the causes is the different types of AIS classes where AIS Class A has a higher sampling rate than AIS Class B. However, this study focus on the early stage in our safety navigation research we have being started. This study examines the computation time required for the ships traffic analysis using the clustering area.

2. Research Method

In this study, a numerical program was developed using the SciLab 6.0.1 program. The numerical programs include decoding AIS data into numerical data, grouping data, and analyzing ship-to-ship traffic status. This research was conducted in the following stages; AIS data decoding, data clustering, and calculation of ship to ship traffic status.

2.1. AIS Data (Automatic Identification System)

This paper is presented in The 4th International Conference on Maritime Education and Training
October 8th, 2020, Makassar, Indonesia

42

AIS system is a ship data communication system that integrates electronic devices such as GPS, RADAR and then transmit data via VHF communications at 161.975Mhz and 162.025MHz frequencies. AIS data consists of static data of ship's attribute (MMSI, ship name, ship dimensions, etc.) and dynamic data including ship's position (lat, lon), speed over ground, and course over ground. The use of AIS on ships was regulated by IMO SOLAS regulations for domestic and international shipping where the ship should transmit AIS data as far as can go and receive data in its covered area. The data used to analyze the traffic environment and navigate the ship in a safe way. VTS receive the AIS data to monitor the ships traffic, guide the navigator, and give a warning for the ship having a potential of sea accident situation.

The process of sending and receive AIS data started with an encoding and decoding process. The process translate the numeric data into character format and vise versa. The AIS data using the NMEA0183 data protocol based on ITU Recommendation M.1371 where AIS data is compiled using the ASCII format regulated in IEC-PAS 61162-100. The data format code starts with the text '!AIVDM' or '!AIVDO' as shown in the following example;

```
!AIVDM,2,1,3,B,55P5TL01VIaAL@7WKO@mBplU@<PDhh000000001S;AJ::4A80?4i@E53,0*3E
!AIVDM,2,2,3,B,1@0000000000000,2*55
```

The AIS device consists of class A and class B AIS receiver-transmitter. The class A provide a better performance than class B, such as, the class A AIS device can transmit every 6 second and class B can transmit data in every 30 seconds.

2.2. Data Clustering

Data clustering is the process of grouping a certain objects based on their characteristics. The duration of the grouping process is directly proportional to the number of criteria and the amount of data to be grouped. In determining the status of ship traffic in a busy channel, the ships can be grouped first based on a certain area. In this study, the data grouping algorithm used is the DBSCAN algorithm [9] with ship-to-ship distance as an expand criteria [5]. Before data clustering using the DBSCAN algorithm, there are two parameters that need to be determined; the safety radius of navigation and the minimum number of objects as initial in calculation. In the ship navigation, the radius is calculated from AIS data. Then, the data clustering algorithm is shown as follows;

Algorithm DBSCAN with expan cluster

```
Input :  D, Eps, MinObj
Output  :  Labels for each sample in D
C = 0
    for each point P in dataset D do
        if P is visited then
           Continue next point
        else
           Mark P as visited
        neighborObj = RegionQuery(P,Eps)
        if sizeof(neighborObj<MinObj then
           Mark P as Noise
        else
        C = next cluster
```

Add P to cluster C
for each point P' in NeighborObj ***do***
   if P' is not visited ***then***
     Mark P' as visited
   NeigborObj' = RegionQuery(P',Eps)
   if sizeof(neigborObk')>= MinObj ***then***
     NeighborObj = NeighborObj joined with NeigborObj'
   if P' is not yet member of any cluster ***then***
     Add P' to cluster C

2.3. Ship to Ship Traffic Status

The traffic status between ships are the state of the ship travels to another ship. Calculation of the traffic status using dynamic AIS data such as; ship position (lat, lon), speed over ground, and course over ground. Based on these dynamic data, the minimum parameter of determining the traffic status of ship to ship were calculated. The parameters were the closest point of approach, the closest distance to point of approach, and the time travel to the closest point of approach.

2.3.1. Collision Prediction Point

The prediction point calculated based on two linear intersection lines. The collision point position (lat_p, lon_p) calculate as follows;

$$dist = 2 \times asin \left(\sqrt{(sin^2(\Delta lat/2) + \cos(lat_o) \times \cos(lat_t) \times sin^2(\Delta lon/2))} \right) \tag{1}$$

$$tta = acos \left((sin(lat_t) - \sin(lat_o) \times \cos(dist)) / (\sin(dist) \times \cos(lat_o)) \right) \tag{2}$$

$$ttb = acos \left((sin(lat_o) - \sin(lat_t) \times \cos(dist)) / (\sin(dist) \times \cos(lat_t)) \right) \tag{3}$$

$$tt12, tt21 \begin{cases} tt12 = tta, \ tt21 = 2 \times \pi - ttb, & \sin(lon_t - lon_o) > 0 \\ tt12 = 2 \times \pi - tta, \ tt21 = ttb \end{cases} \tag{4}$$

$$\propto_1 = tt13 - tt12 \tag{5}$$
$$\propto_2 = tt21 - tt23 \tag{6}$$
$$\propto_3 = acos \left(-\cos(\propto_1) \times \cos(\propto_2) + \sin(\propto_1) \times \sin(\propto_2) \times \cos(dist) \right) \tag{7}$$
$$d_{13} = atan \left(\frac{\sin(dist) \times \sin(\propto_1) \times \sin(\propto_2)}{\cos(\propto_2) + \cos(\propto_1) \times \cos(\propto_3)} \right) \tag{8}$$
$$lat_p = asin \left(\sin(lat_1) \times \cos(d13) + \cos(lat_1) \times \sin(d13) \times \cos(tt13) \right) \tag{9}$$
$$\Delta lon13 = atan \left(\frac{\sin(tt13) \times \sin(d13) \times \cos(lat_o)}{\cos(d13) - \sin(lat_o) \times \sin(lat_p)} \right) \tag{10}$$
$$lon_p = lon_o + \Delta lon13 \tag{11}$$

2.3.2. Calculation of CPA, DCPA and TCPA

The actual value of CPA is determined by the actual distance between two ships involved in an encounter situation and identified limit values are needed to take a warning. The distance and bearing were calculated from the V_t (target vehicle) and V_o (own vehicle) as follows;

This paper is presented in The 4th International Conference on Maritime Education and Training
October 8th, 2020, Makassar, Indonesia

44

$$Dif_{long} = Long_t - long_o \tag{12}$$
$$Dif_{lat} = Lat_t - Lat_o \tag{13}$$
$$MP = 7915.7 \, log \left[tan \left(\frac{pi}{4} + \frac{Lat}{2} \right) \left(\frac{1-esin(lat)}{1+esin(lat)} \right)^{e/2} \right] \tag{14}$$
$$DMP = MP_t - MP_o$$
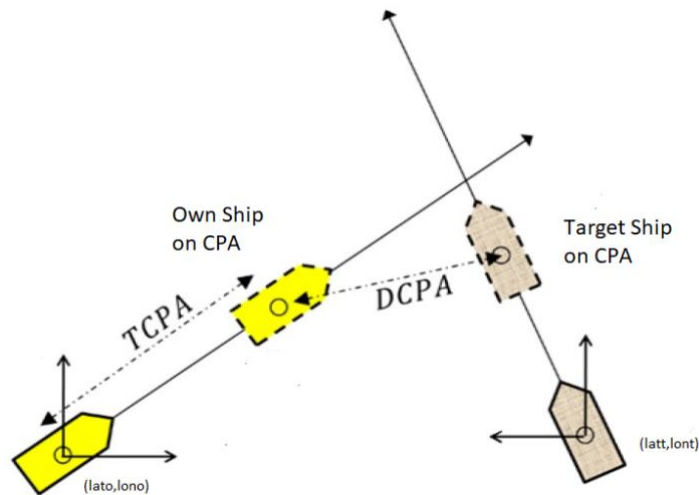$$Bearing = arctan(DMP/Dif_{Long}) \tag{15}$$
$$Dist = Dif_{Lat}/\cos(Bearing) \tag{16}$$

Figure 1. CPA, DCPA, and TCPA in the crossing ship situation

The distance between the two ships's position in the ship traffic can change to be closer, constant or take away. These changes are identified by calculating the change in the distance of the previous position with the current position. The negative change will indicate the distance of the two positions are close to and the positive change indicates the distance of the two positions are get away. In the Figure 2 the change of the distance of two ships illustrated from time to time. The distance of the ship A and ship B from the time t_0 to t_1 is going closer and the distance after from t_1 to t_2 is going to take away. The CPA of the two ships should be considered when the distance is in the radius of the identified limit value.
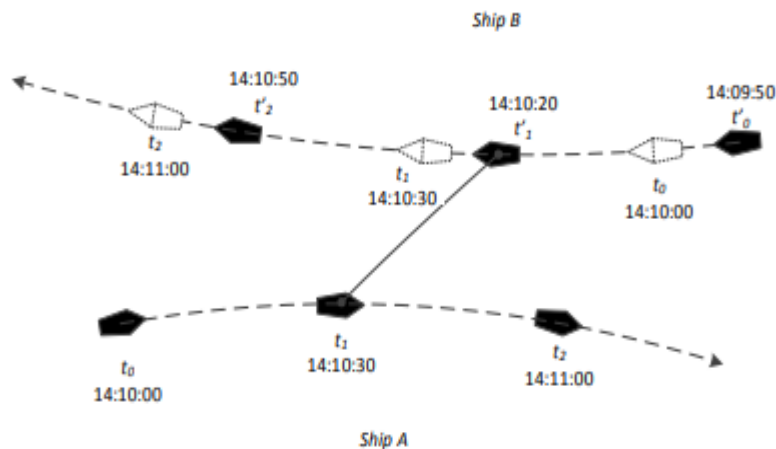
Figure 2: Illustration of the changes distance of two ships moving

The relative speed (Sr) and relative course (Cr) are calculated as shown below;

$$S_r = \sqrt{Sog_o^2 + Sog_t^2 + 2Sog_o Sog_t \cos(Cog_t - Cog_o)}$$
(17)

$$C_r = \begin{cases} Cog_o - arcos\left(\frac{S_r^2 + Sog_o^2 - Sog_t^2}{2S_r Sog_o}\right) & Cog_o < Cog_t \\ Cog_o + arcos\left(\frac{S_r^2 + Sog_o^2 - Sog_t^2}{2S_r Sog_o}\right) & Cog_o \geq Cog_t \end{cases}$$
(18)

Then, DCPA and TCPA were calculated using the following equation

$$DCPA = dist\left(sin(C_r - Cog_o - Bearing - \pi)\right)$$
(19)

$$TCPA = dist\left(sin(C_r - Cog_o - Bearing - \pi)/S_r\right)$$
(20)

3. Results and Analysis

In this research, a numerical program has been developed that can be used to translate AIS data in the form of text or characters into numerical data that can be used to analyze ship traffic. Simulation of time computation of ship traffic analysis for number of ships 50, 100, 200, 300, 400, and 500 ships and the number of clusters 1, 2, 4, 6, and 9 clusters were also conducted. The ship's position generated in the area of 30.91 km square or 0.0025 deg square. The ship's state generated were position (lat, lon), speed over ground, and course over ground randomly. This data assumed as AIS data received by VTS. The simulation calculated the intersection point of collision, distance between ships, bearing angle, CPA, TCPA, and DCPA. The simulation conducted using a laptop with Intel Core i5 6200U processor and the speed clock is 2.1GHz. The number of ships and clusters in this analysis is to describe the density of the shipping area and the number of ship traffic areas monitored by VTS.

Using the DBSCAN algorithm, all ship's position against other ships were analyzed based on the ship-to-ship distance criteria for each traffic in the cluster area. When a ship's position was within the radius of the ring guard or intersects with another ship's ring guard, the ship's position is marked with a '+' as shown on the Figure 3. The position of the ship which is marked with '*' and not close to the position of the ship with '+' indicates that the ship is in a safe condition towards other ships. The position of the ship which is in around the sign "+" the traffic status is analyzed by calculating the intersection point, CPA, DCPA, and TCPA.
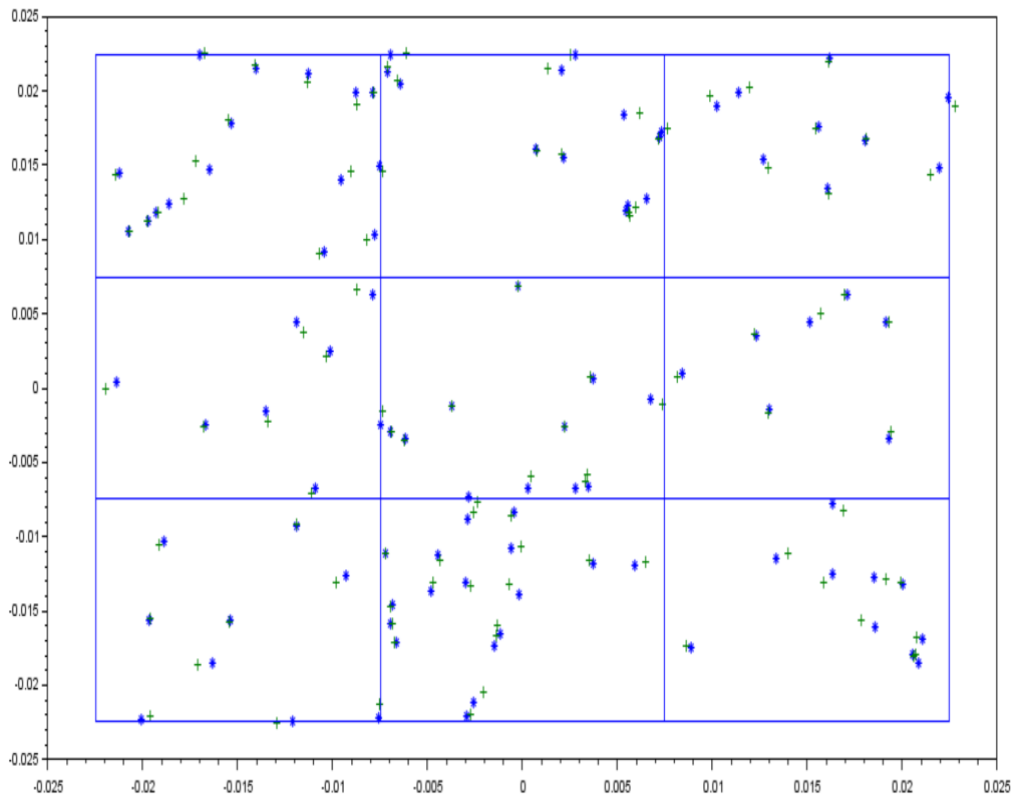
Figure 3 Simulation of 200 ships with 9 data clusters for ship-ship traffic analysis

The results of the computation time for the variation in the number of ships and the number of clusters show that the more the number of clusters, the faster the calculation time will be, as seen in Figure 4 and Table 1. The increase in computation time for all the number of clusters follows a second order non-linear trend. Likewise, the average calculation time for each ship requires a cycle time of under 0.2 seconds for the number of ships 500 and 9 clusters as shown in Table 2.

Table 1 Computation time of numeric analysis of ship traffic for number of ships and number of clustering

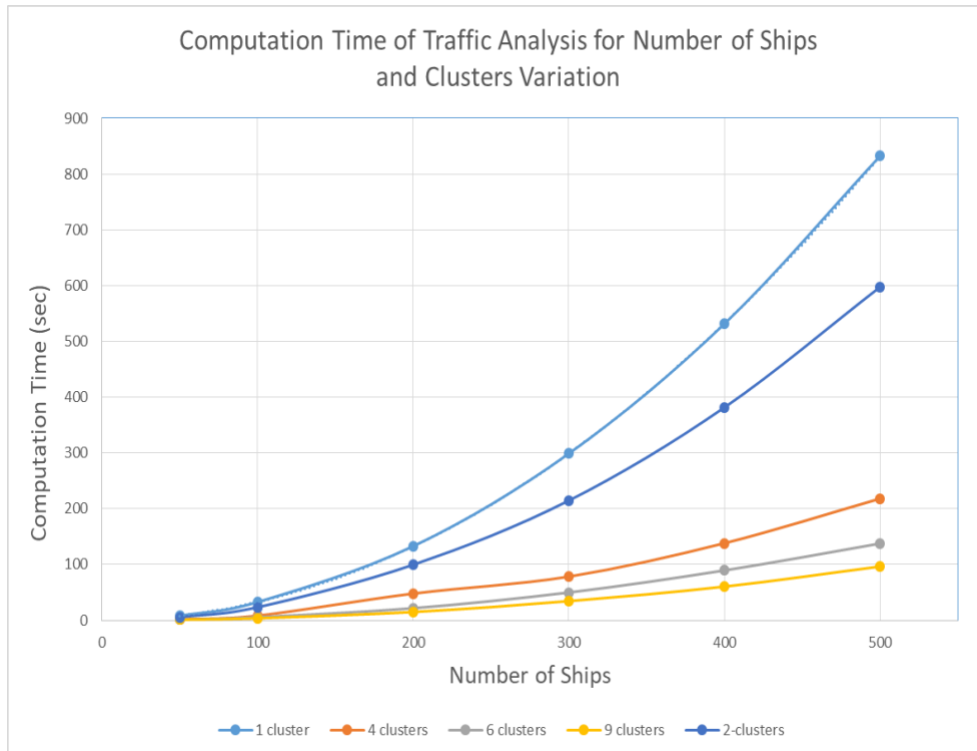| Number of Ships	Number of Clusters				
	1	2	4	6	9
50	8.13	5.75	1.92	1.27	1.02
100	32.83	23.51	8.38	5.5	3.81
200	132.9	99.63	47.93	21.83	14.96
300	298.91	214.66	78.36	49.89	34.49
400	531.42	381.50	138.27	89.46	60.47
500	832.21	597.44	218.13	137.72	96.47

Figure 4 Simulation of time computing for ship traffic analysis with number of ships and number of clusters variations

Table 2 Mean of Computation time for numeric analysis of ship traffic, with number of ships and number of clustering variations

Number of Ships	Number of Clusters				
	1	2	4	6	9
50	0.163	0.115	0.038	0.025	0.020
100	0.328	0.235	0.084	0.055	0.038
200	0.665	0.498	0.240	0.109	0.075
300	0.996	0.716	0.261	0.166	0.115
400	1.329	0.954	0.346	0.224	0.151
500	1.664	1.195	0.436	0.275	0.193

If we look at the reduction in computation time for the ship traffic analysis at each additional number of clusters, it shows that in the four clusters of ship traffic analysis, the largest reduction in computation time is obtained when compared to one cluster, although overall the fastest computation time is in the nine clusters, as shown in the Figure 5 and the data showed in the Table 3.

This paper is presented in The 4th International Conference on Maritime Education and Training
October 8th, 2020, Makassar, Indonesia

48

Figure 5. Reduction in Computation time for ship traffic analysis for every clusters and number of ships comparted to the computation time required for single cluster.

Table 3 Reduction in time computation for number of ships and number of ship's clusters

Number of Ship's Cluster	Number of Ships					
	50	100	200	300	400	500
	Reference of Computation Time for 1 Cluster					
	8.13	32.83	132.9	298.91	531.42	832.21
1	0	0	0	0	0	0
2	-2.38	-9.32	-33.27	-84.25	-149.92	-234.77
4	-3.83	-15.13	-51.70	-136.30	-243.23	-379.31
6	-0.65	-2.88	-26.1	-28.47	-48.81	-80.41
9	-0.25	-1.69	-6.87	-15.4	-28.99	-41.25

4. Conclusion

The computation time analysis illustrates the time required by the computation system on VTS to analyze ship traffic within the VTS coverage area or in certain shipping areas, especially in areas with a high ship traffic density. If the computation time in ship traffic analysis is long enough, it means that there is a loss of time spots in monitoring ship traffic. This can have an impact on the possibility of sea accidents that cannot be monitored or the action of monitoring the ship traffic are made late. To prevent the impact, the computation time of VTS should be short as possible.

Based on the simulation results, the computation time increases quadratically to the amount of data. Using the cluster area method and the DBSCAN algorithm has shown a significant reduction in computation time. Thus, VTS operators can consider the clustering area to increase the speed of ship traffic analysis based on the number of ships in covered area.

This paper is presented in The 4th International Conference on Maritime Education and Training
October 8th, 2020, Makassar, Indonesia

49

Acknowledgement

References

[1] A. Lazarowska, "Ship's trajectory planning for collision avoidance at sea based on Ant Colony Optimisation," *J. Navig.*, vol. 68, no. 2, pp. 291–307, 2015.

[2] L. P. Perera, P. Oliveira, and C. Guedes Soares, "Maritime Traffic Monitoring Based on Vessel Detection, Tracking, State Estimation, and Trajectory Prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1188–1200, 2012.

[3] E. Tu, G. Zhang, L. Rachmawati, E. Rajabally, and G. Bin Huang, "Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey From Data to Methodology," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–24, 2017.

[4] D. Gattuso, S. Napoli, and A. G. Meduri, "Safety of the navigation in congested maritime area: The case of the Messina Strait," *Eur. Transp.*, vol. 42, pp. 83–101, 2009.

[5] R. Zhen, M. Riveiro, and Y. Jin, "A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traf fi c surveillance," *Ocean Eng.*, no. February, pp. 1–10, 2017.

[6] H. Hilgert and M. Baldauf, "A common risk model for the assessment of encounter situations on board ships," *Dtsch. Hydrogr. Zeitschrift*, vol. 49, no. 4, pp. 531–542, 1997.

[7] H. J. Lee and K. P. Rhee, "Development of collision avoidance system by using expert system and search algorithm," *Int. Shipbuild. Prog.*, vol. 48, no. 3, pp. 197–210, 2001.

[8] S. L. Kao, K. T. Lee, K. Y. Chang, and M. Der Ko, "A fuzzy logic method for collision avoidance in vessel traffic service," *J. Navig.*, vol. 60, no. 1, pp. 17–31, 2007.

[9] X. Xu, H. Peter kriegel, M. Ester, and J. Sander, "Density-Based Clustering in Spatial Databases : The Algorithm GDBSCAN and Its Applications," vol. 194, pp. 169–194, 1998.

This paper is presented in The 4ᵗʰ International Conference on Maritime Education and Training
October 8ᵗʰ, 2020, Makassar, Indonesia

50