Data Mining Analysis on Ships Collision Risk and Marine Traffic Characteristic of Port Klang Malaysia Waterways from Automatic Identification System (AIS) Data

Conference Paper · March 2019

CITATIONS
0

READS
21

10 authors, including:

Masnawi Mustaffa
Universiti Teknologi MARA
21 PUBLICATIONS **52** CITATIONS
SEE PROFILE

Shaharudin Ahmad
Universiti Teknologi MARA
10 PUBLICATIONS **7** CITATIONS
SEE PROFILE

Ab Malik Marwan Ali
Universiti Teknologi MARA
138 PUBLICATIONS **841** CITATIONS
SEE PROFILE

Nadim Ahmad
Universiti Teknologi MARA
2 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Data Mining Analysis on Ships Collision Risk and Marine Traffic Characteristic of Port Klang Malaysia Waterways from Automatic Identification System (AIS) Data

M. Mustaffa, S. Ahmad, M.I.H.M. Nasrudin, K.A. Sekak, N.A. Aini, A.M.M.Ali, M.F.M. Taib, M.Z.A.Yahya, O.H.Hassan, A.K. Razali, M.H.M.Jais

Abstract— **Port Klang is one of the busiest ports in the worlds and have played an important role to ensure the import and export activities towards Malaysian economy sector. Port Klang located in a busy marine sea route of Strait of Malacca received a great amount of vessels making it a high density port and more likely exposed to the collision risk. From the analysis of Automatic Identification System (AIS) data, the marine traffic pattern of the study area will be characterized. The raw AIS data collected through AIS receiver that was installed at Universiti Teknologi Mara, Shah Alam, Malaysia. The data was collected from July 2016 until December 2016. A web-based application was developed to decode the raw data collected for analysis and characteristic purpose. The decoded raw data were presented statically and through graphical method. The variable that need to be consider such as number of vessels, type of vessels, course over ground, and speed over ground of the vessels. The path for dangerous category vessels was also plotted.**

Index Terms— **AIS data, big data analysis, Data mining, ship collision, traffics density, ship navigation.**

I. Introduction

Traffic density is an approach to studying the traffic flow and predicts its behavior in future time horizons. The definition of traffic density is relative to the traffics existence domain in the space or area as a time-varying high-level view of travel [11].

Malaysia is a country that surrounded by sea. Thus, its economic trading activities have been depending on its ports and shipping activities. Both ports and shipping are crucial aspect to enhance Malaysian trade as the result, the economic prosperity rise yearly [1]. Due to its location that located in Straits of Malacca, Malaysia is one of the busiest shipping waterways. Traffic pattern and ship collision risk is closely related in maritime study. Both aspects is important so that the safety aspect of marine navigation is in the highest level.

There are rapid increase in data receive each day as Port Klang continue becoming one of the most important trading hub in Malaysia. However, recent study show that with the North Sea route opening, the traffic pattern at Port Klang will be affected so that the ship collision risk is reduced [7, 8, 10]. The finding of the study shows that there will always changes in the traffic pattern.

I. Automatic Identification System (AIS) Background

AIS is system used on ships and by vessel traffic services (VTS) for tracking purposes. AIS contain a lot of data that have huge potential to be used as information in term of marine navigational safety. However there are few system and method to analyze it to be transform to more useful data for the navigator. The marine traffic pattern of the study area is needed to be characterized [1].

Due to the high number of traffic density, adding with geographically narrow strait, marine navigational hazard will emerge. Thus, an organized monitoring system of vessels traffic is an important aspect in marine navigation. One of the main focuses of application is to provide safe navigation of strait waterways [6]. Collision can be prevent by appropriate training of officer in charge to analyze AIS data. [4, 5].

AIS transponder is a compulsory equipment for each all larger sea-going vessels (>300 GT), all passenger vessel and new ship. AIS transponder need to be installed on board before allowed sailed on the sea [3]. The International Maritime Organization instructed the installation in order to ensure a safe navigation among ships [6].

AIS are designed to provide information about the ship to other ships and to coastal authorities automatically for ship 300 gross tonnage and above engaged on international voyages, cargo ships of 500 gross tonnage and upwards not engaged on international voyages and all passenger ships irrespective of size[2].

Characteristic of marine traffic pattern and density are more likely harder to be characterized. A study suggest a visualization model using Automatic Identification System

Manuscript received December 7, 2018. This work was supported by the Ministry of Higher Education, Malaysia under Fundamental Research Grant Scheme (FRGS/1/2017/STG02/UITM/02/4).

M. Mustaffa is with Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia (Email: masnawimustaffa@gmail.com).

(AIS) to estimate the ship's real-time maritime traffic situation [9]

II. Methodology

The raw data collected for this study is obtained from the AIS receiver which was placed on top of 5th floor building in our campus, Malaysia, at altitude of 31 meter and facing Port Klang waterways and Malacca Strait for a better coverage.

Fig 1. Equipment Setup [8]

The process of analyzing the raw has been developed in web-based environment and as a one-stop web-based monitoring system [12,13]. This analyzing of raw data is to allow the loading of AIS files in its natural format which are process as discrete bundle of information. The AIS data collected by the receiver is in ASCII (American Standard Code for Information Interchange) data packet.

III. Results and Discussions

A. Daily Data

In general, the number of vessels trend is differ and unique from each month as shown in Fig 2. Besides, there are few unavailable set of data for several days in August, September, October, November and December. In July, the traffic pattern showing a clustered trend as shown in Fig 3. There are no particular trends showed by the number of vessels with total of 6731 vessels recorded. The highest number recorded is 275 vessels in 8/07/2016 and 6/07/2016 record the lowest number with only 107 vessels. The average number is 212 vessels. The total number of vessels recorded for July is 6580 vessels.

For August as shown in Fig 4, the number of vessels is fluctuating each day with trend packing closely to the median of the graph. The median for August is 189 vessels. The highest number recorded is 263 on 16/08/2016 and 113 is the lowest on 21/08/2016 and 23/08/2016. The total number of vessels recorded in August is 5693 vessels. There are no data recorded on 20/08/2016 and 26/08/2016.

For September (Fig. 5), the number of vessels can be considered constant with ranging from 275 to 110 with no drastic change in the value except from 1st to 3rd of September 2016 where there are no data recorded. The anomaly in the trend might also due to the error. The median for September is 211 vessels. The highest number recorded is 275 on 13/09/2016 and 110 is the lowest on 28/09/2016. The total number of vessels recorded is 5699 vessels.

For October (Fig. 6), the median for is 202 vessels. The highest number recorded is 255 on 7/10/2016 and 94 is the lowest on 11/10/2016. Logically, the number is unreliable since it is far off compared to the median value. The zero value from the graph on 12th, 15th, 16th, 17th, 25th, 26th, and 27th of October is missing from the AIS database. There are high likely there are certain number of vessels on the date. The total number of vessels recorded for October is 4846 vessels

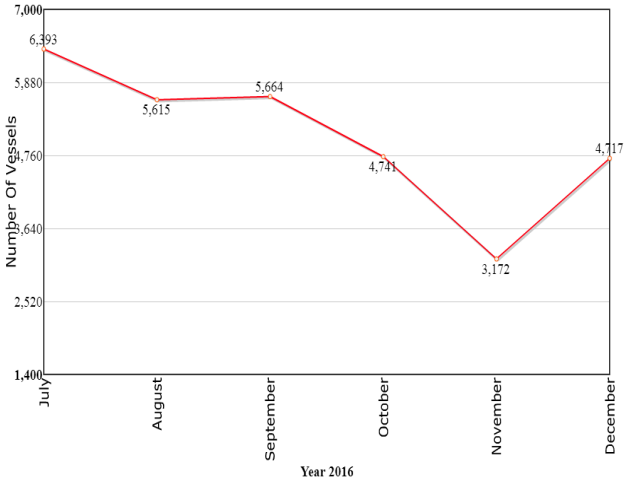

Fig 2. Number of recorded ships from July 2016 to December 2016

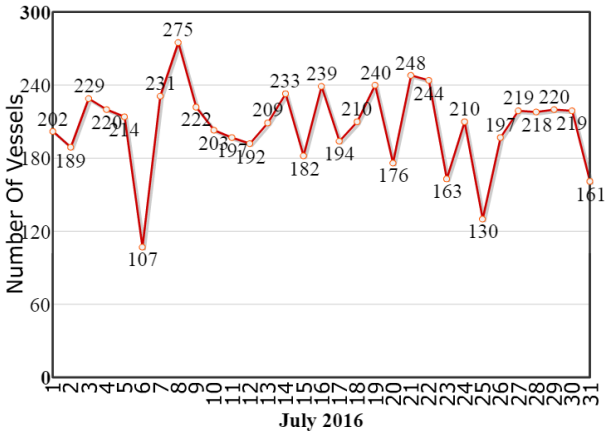

Fig 3. Marine traffic statistic for July 2016

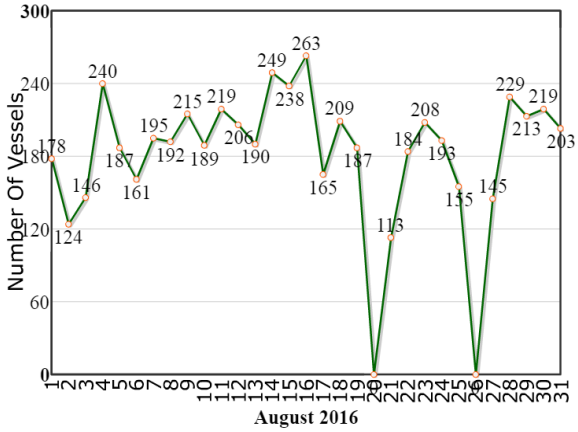

Fig 4. Marine traffic statistic for August 2016

The number of vessels generally increasing each day with a decreasing start until 5/11/2016 before increasing generally until the end of November as shown in Fig 7. The median for is 198 vessels. The total vessel recorded is 5141. The highest number recorded is 294 on 2911/2016 and 78 is the lowest on 20/11/2016. Logically, the number is unreliable since it is far off compared to the median value. The zero value from the graph on 11th, 12th, 13th, and 27th of November is missing from the AIS database. There are high likely there are certain number of vessels on that date. There might be due to some errors occurs during data collection period

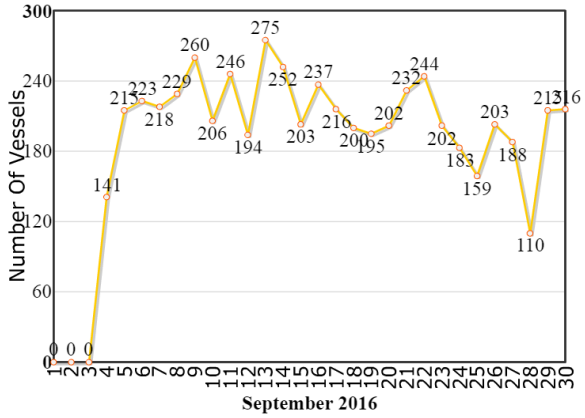

Fig 5. Marine traffic statistic for September 2016

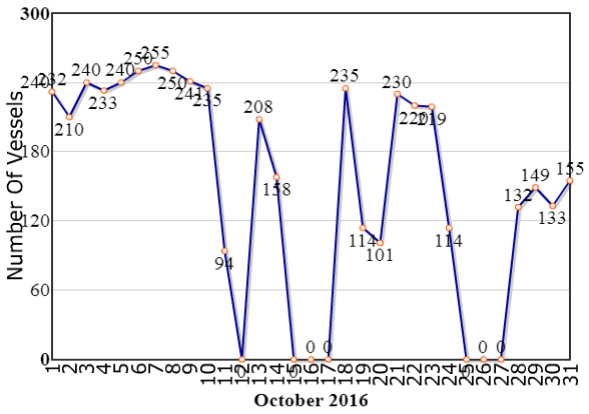

Fig 6. Marine traffic statistic for October 2016

The number of vessels increasing each day up until 06/12/2016 where it reach the highest number of vessels recorded of 315 vessels in the month of December with as shown in Fig 8. Then there are no AIS data received by the data receiver until 13th December. The median is 192 vessels. The lowest number recorded is 108 on 31/12/2016. The total number of vessels recorded is 4799 vessels.

Fig 9 shows the statistical characteristic of number of vessels that use the Port Klang coastal area waterways throughout the study period. From the graph generally, the highest number of vessels recorded was on the middle of the month. In the beginning and the end of the month, the pattern is not constant for each month. There are big different each month especially on the end of the study period. For example November record its second highest while December record its lowest number of vessels.

The daily heaviest traffic for each month have been recognized. The data was collected from 8th July, 16th August, 13th September, 7th October, 29th November and 6th December as the date give the highest traffic density.

B. *Hourly Data*

The pattern and statistical characteristic for number of vessels in hourly have been plotted as shown in Fig. 10. It divided into three main groups. The first group is morning starting from 1 a.m.to 2 a.m., then 9 a.m. to 11 a.m. for the day groups. For evening group, 4 p.m. to 5 p.m. recorded the highest traffic density.

Risks of collision are more likely to occur during this range of time stamp. However, the traffic flow is higher in the morning group where there are big numbers of vessels going in and out of the waterways thus increasing the chance of collision. On the sample from 8th July, there are 31 vessels going in and out of the study location, the highest for the morning group. The biggest number of vessels going in and out was recorded on 16th August with 49 vessels. This indicates that the highest traffic density doesn't necessarily give the higher risk of collision.

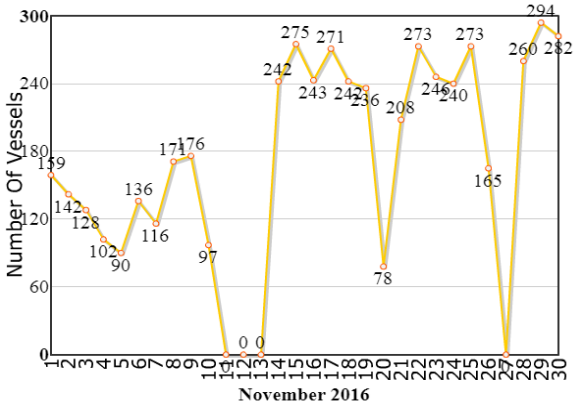

Fig 7. Marine traffic statistic November 2016

In the evening group, the biggest different was 76 vessels making it having the highest risk of collision in term of number of vessels in the time stamp. The data was collected on 13th of September. In the other sample, the difference are ranging only from 3 vessels to 10 vessels maximum making it low in risk collision

C. *Type of Vessels Data*

Type of vessels in the Port Klang waterway area was determined and recorded. There are more than 25 different type of vessels manage to be recorded. From the total, few types was sorted out from the consideration of the study as their frequency using the Port Klang coastal area waterways are too small. In the day time, vessels such as fishing vessels, law enforcer and pleasure craft only recorded in range of 0 to 5. Thus, the risk of collision for these types of vessels is lower compared to other type of vessels.

Vessels such as carrier, cargo, container, tanker, tug, vehicles and passenger carrier and others was identified as the dangerous type of vessels. Fig 11 shows the number of vessels for different type of vessels for the highest density traffic day in each month through the study period. The type of vessels from 8th July, 16th August, 13th September, 7th October, 29th November and 6th December was used.

ISBN: 978-988-14048-5-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Container and tanker is the most frequent vessels that travel using Port Klang coastal area. Container and tanker is the most frequent vessels that travel using Port Klang coastal area.

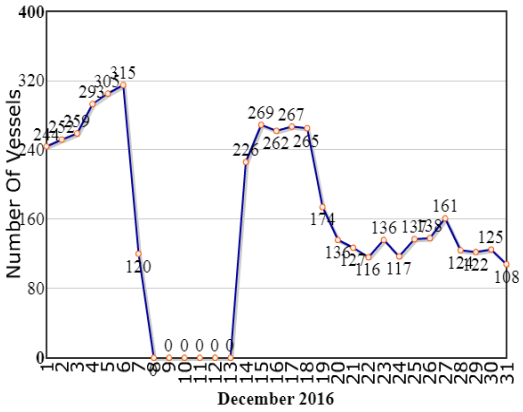

Fig 8. Marine traffic statistic for December 2016

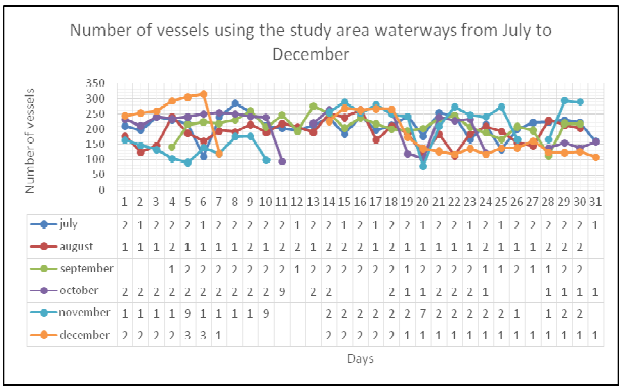

Fig 9. Number of vessels recorded in the study area from July to December 2016

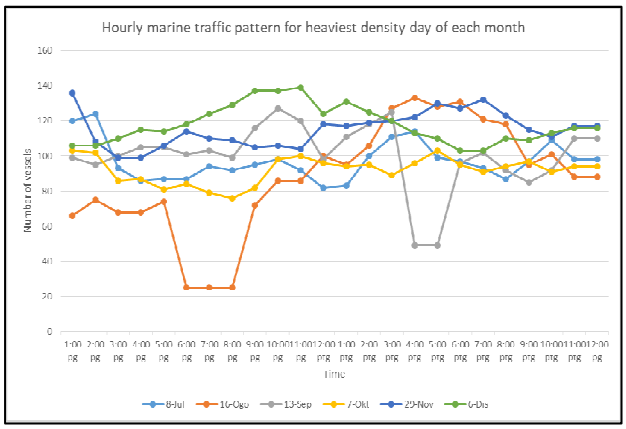

Fig 10. Density of vessels by hour

Tanker and container generally become the most frequent type of vessels using the study area waterways. From the sample taken, on 8th July (31.21%), 16th August (26.86%), 29th November (30.4%), and 6th December (31.07%), tanker is the highest number of vessels recorded. While the rest, 13th September (37.73%) and 7th October (30.04%) recorded container as their highest number of vessels using the waterways.

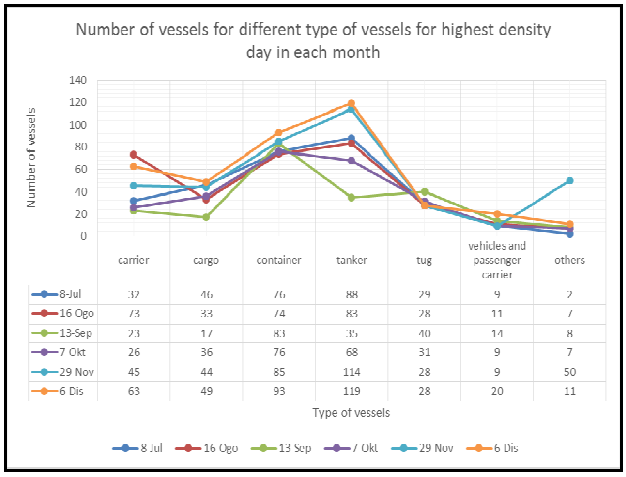

Fig 11. Number of vessels for different type of vessels for highest density day in each month

D. *Vessels Trajectory*

The general trajectory for dangerous category vessels such container, and tanker was plotted on the satellite image (Fig. 12) . Form the plotting of a sample from 8th of July 2016, a general pathway of those type of vessels can be determined. For the risk of ship collision, there are high likely of occurrence because there are few data transmitted when the vessels approaching Port Klang coastal area. The AIS data received delayed for about 73 km from the last it was transmitted. Although it was further from the Port Klang coastal area where the density is higher, more packed and congested, the delay of information is dangerous as the information of the surrounding vessels is unknown. This make the decision making for a safer route that should be taken can't be made in an appropriate time range until the danger is too near. This make the risk of collision are more likely to be happened in this type of scenario.

Fig 12. Density of vessels trajectory

ISBN: 978-988-14048-5-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

IV. Conclusion

This study shows that the number of vessels using the study area waterways ranging from 78 vessels on 20th November to the highest recorded on 6th December with 315 vessels.

The anomaly of data such as the big drop of number of vessels from day to the next day (6th to 7th December), retrieve data for a few days, and big delayed of AIS data transmitted by the vessels affect the study and data analysis process as the exact value can't be determined.. The result of the current study are necessary to characterize the Port Klang coastal area. The risk of collision can be determined by taking a suitable precaution steps when the trend of high density traffic pattern scenario emerge. So that the safety of the waterways of the study area can be improved as it play an important role for nation economic trading.

Acknowledgment

This research was supported by Fundamental Research Grant Scheme (FRGS) Grant No. **FRGS/1/2017/STG02/UITM/02/4**. Author would like to thank Office of Research Management Centre (RMC), UiTM and Ministry of Higher Education for sponsoring the research.

References

[1] Harati-Mokhtari, A., Wall, A., Brooks, P., & Wang, J., "Automatic Identification System (AIS): Data Reliability and Human Error Implications," Journal of Navigation, vol. 60, no. 3, pp. 373-389, Sep 2007.

[2] Jun Min Mou, Cees van der Tak, Han Ligteringen, "Study on Collision Avoidance In Busy Waterways By Using AIS data," Ocean Engineering, ol. 37, no.5–6, pp. 483-490, April 2010.

[3] PAN Jiacai, JIANG Qingshan, HU Jinxing, SHAO Zheping, "An AIS data Visualization Model for Assessing Maritime Traffic Situation and its Applications", Procedia Engineering, vol. 29, pp. 365-369, 2012.

[4] Marine Accident Investigation Branch, "Report on the investigation of the collision between ACX Hibiscus and Hyundai Discovery in the approaches to the eastern Singapore Strait TSS at 0756 local time on 11 December 2011", Southampton,U .K.: MAIB, pp. 39., June 2013.

[5] Muhammad Badrus Zamana, Eiichi Kobayashib, Nobukazu Wakabayashic Adi Maimun, "Risk of Navigation for Marine Traffic in the Malacca Strait using AIS", Procedia Earth and Planetary Science, vol. 14, pp. 33-40, 2015

[6] N.S.F Abdul Rahman , A.H. Saharuddin, R. Rasdi, "Effect Of The Northen Sea Route Opening To The Shipping Activities At Malacca Straits", International Journal of e-Navigation and Maritime Economy, vol. 1, pp. 85-98, Dec. 2014

[7] Hendra Saputra, Mufti Fathonah Muvariz, Sapto Wiratno, Satoto, Jaswar Kotob "Estimation Of Exhaust Ship Emission From Marine Traffic In The Straits Of Singapore And Batam Waterways Using Automatic Identification (AIS) Data", Jurnal Teknologi, vol. 77, no. 23, pp. 47-53, 2015

[8] Masnawi Mustaffa, Munawwarah Abas, Shaharudin Ahmad, Nazli Ahmad Aini, Wan Faezah Abbas, Syahrul Afzal Che Abdullah, Nur Idora Abd. Razak, Mohamad Yusof Darus,"Marine Traffic Density Over Port Klang Malaysia Using Statistical Analysis Of AIS Data: A Preliminary Study", Journal ETA Maritime Science, 4(4): 333-341, Nov. 2016

[9] Cheng Tsou, "Online Analysis Process On Automatic Identification System Data Warehouse For Application In Vessel Traffic Service", J Engineering for the Maritime Environment, vol. 230, no. 1, pp. 199-215, Feb. 2016

[10] M. Mustaffa, S Ahmad, N H Ahmat, "Mapping Vessel Path of Marine Traffic Density of Port Klang, Malaysia using Automatic Identification System (AIS) Data", International Journal of Science and Research (IJSR), 4(11):245-248, Nov. 2015

[11] Masnawi Mustaffa, Shaharudin Ahmad, Radhiah Mohd Nor, Nur Hidayah Binti Zahari, "Preliminary Study on Air Traffic Density of Peninsular Malaysia using Visual Flight Path Trajectories from Automatic Dependent Surveillance-Broadcast (ADS-B) Data", International Journal of Engineering & Technology (IJET), Vol 7, No 4.25, 2018

[12] Wan Faezah Abbas, Masnawi Mustaffa, Sulaiman Shaari, Ahmad Maliki Omar , Helme Mohd Yusop, "Development of Photovoltaic Monitoring System Web Portal", International Conference In Software Engineering and Computer Systems 2009 (ICSECS 2009), 186-191, Jan. 2009

[13] M. Mustaffa, W.F. Abbas, P. Mazlan, R.A.H.A. Kadir, L. Rimon, A.R. Ghazali, "The Photovoltaic Monitoring Centre Portal: From Inception to Development and Maintenance", 9th International Seminar in Sustainable Environment and Architecture (SENVAR2008), Jan. 2008