Human-Machine Interface Evaluation in Engine Supervisory Control through Alarm Performance Assessment

Adi Mas Nizar*, Takashi Miwa, Makoto Uchida

Graduate School of Maritime Sciences, Kobe University, Japan
* Corresponding Author. Email: 185w604w@stu.kobe-u.ac.jp

Synopsis

An alarm system installed onboard a ship is often reported as a nuisance. It was even blamed as one of the causes of the accident. These facts draw the opposite function of the alarm system as the first layer of support to the operators during troubleshooting in the engine supervisory control. Additionally, the number of installed sensors on the machinery system increases with the automation's implementation, likewise the number of alarms stored on the engine control consoles. Therefore, alarm management, a common practice in the onshore industry, becomes coherent also in maritime operations. It is beneficial for the operators who handle the alarms directly, and also for the onshore support staff, to grasp the onboard work situation. As the initial step of the alarm management, we conducted the alarm performance assessment with actual alarm data from the ocean-going vessel. The alarm data was retrieved in a data set containing the alarm name, alarm activation time, alarm deactivation time, and sensor reading value. This study developed several performance indexes based on modified methods ready on the literature: chattering index to categorize the nuisance alarms, similarity index between unique alarms, and similarity index between alarm floods. The actual data ship analysis shows the alarm performance assessment was able to discover several nuisance alarms and alarm floods. Thereafter, the practices in alarm management can be considered to minimize or eliminate these alarms, such as reconfiguring the alarm set-point, applying a delay time, and preparing the response strategy for alarm floods. Although alarm performance assessment only evaluated a small part of the human-machine interface, it provides added value in the age of digitalization and massive data communication. Alarm performance assessment can be a consideration for both onboard operators and shore management to maintain safe operations.

Keywords: Alarms, Engine Control Console, Human-Machine Interface

1. Introduction

Implementation of automation brings more sensors installed in the engine room. It aligns with the increasing number of alarms stored in the engine control console that are ready to be announced. Along with the automation, the onboard engine department's workload is increase with the increment of information used in the process control (Man *et al.* 2018). This condition is also elevated since the stress induced by the increasing number of alarms onboard (Lundh *et al.* 2011). From another finding, only a small group of unique alarms generates a significant number of annunciated alarms (Rødseth *et al.* 2006).

The alarms system onboard a ship is often reported by seafarers as a distraction since there are nuisance alarms, excessive workload made by multiple alarms, and alarm messages that are sometimes confusing or unclear (Jones *et al.* 2006). A high frequency of false alarms negatively affects the operators as they lose their trust in the alarm system, delay the response, and even neglect it. Another issue is made by the alarm flood, where the operators were overwhelmed by a continuously or long sequence of alarms and made the operators hardly take a response action. The object lesson comes from the report of water flooding at Emma Maersk's engine room on February 1st, 2013 (DMAIB 2013). The incoming water that initially came from the shaft tunnel was caused by mechanical breakdown. However, the condition got even worse because the alarm flood constantly disturbed the operator in troubleshooting action. It is essential to prevent alarm flooding during a major incident and provide the operators with clear and prioritized alarms, allowing them to perform efficient response actions (Chilcott and Kennedy 2018).

Such alarm system issues also exist in the onshore process control, and it has been countered by conducting an alarm performance assessment. This practice is guided by ANSI/ISA 18.2 (ANSI/ISA 2016) and EEMUA 191 (EEMUA 2009) to ensure alarm system management provides adequate support to the operators. In the maritime area, such guidance was included in the Code on Alerts and Indicators 2009 (IMO 2009) and more extensively with bridge alarm management (IMO 2010). Most of the guidelines argue, for instance, to prioritize the alarms based on their indication of severity, function, and allowable response time. Lowest priority alarms

Author's Biography
Adi Mas Nizar is Ph.D candidate in Marine Engineering, Kobe University. He hold the bachelor degree from ITS Surabaya, and master degree from Kobe University in Marine Engineering field. His main research interests concern the human factors in the maritime operation.
Takashi Miwa is Associate Professor, Graduate School of Maritime Sciences Kobe University. He is in charge of Engine simulator training for marine engineering course students in faculty. He has 15 years of experience in marine engineering. Today, he is studying the effective resource management of the engine room.
Makoto Uchida is a Professor at the Graduate School of Maritime Sciences, Kobe University. He is working to improve the ship propulsion performance and the function of marine engine operation management.

 https://doi.org/10.24868/10712

have the least severe consequences and longest allowable response time, and highest priority alarms have the most severe consequences and allowable shore response. The guidelines also suggest the adjustment of the alarms ratio into three levels: 80% low level alarms, 15% medium level alarms, and 5% high level alarms.

The objective of a human factor study is to guarantee the system is optimized by considering the limitation of the human operator. Alarm management corroborates this objection by ensuring the alarm system effectively supports the operator. To promotes the alarm performance assessment as the initial step of alarm management, this study adopted several performance indexes and demonstrated it with the actual data from the ocean-going vessel.

2. Method and Analysis

We retrieved alarm data from an actual ocean-going vessel from a shipping company. As shown in Table 1, the data structure consists of the time when the alarm is activated or deactivated, alarm id, alarm description, alarm name, process value, and event (alarm or recovery). The classification of alarm tag that consists of alarm id and alarm name is considered to cover two different alarm thresholds. For instance, an alarm with ID 1919 was logged with a HIGH and LOW threshold. It is a necessity to assign a separate TAG1919.LOW and TAG1919.HIGH. The alarm tag is also used to hide the shipping company's confidential data in this body text.

Table 1 Data Structure

Time	ID	Description	Name	Value	Event	Tag
2020/06/18 0:16:46	904	WECS COMMON FAILURE	FAILURE		ALARM	TAG0904.FAILURE
2020/06/18 0:17:08	914	M/E CONTROL OIL PUMP #1	NORMAL		RECOVERY	TAG0914.NORMAL
2020/06/18 0:17:08	612	M/E #2 PISTON COOL LO TEMP	NORMAL	54^0C	RECOVERY	TAG0612.NORMAL
2020/06/18 0:17:08	904	WECS COMMON FAILURE	NORMAL		RECOVERY	TAG0904.NORMAL
2020/06/18 1:41:33	200	#1 MAIN AIR RESERVOIR PRESS	LOW	1.63MPA	ALARM	TAG2001.LOW

2.1. Chattering Alarm

A chattering alarm is an alarm tag with a short interval between activated and deactivated times (ANSI/ISA 2016). Rule of thumb defines the alarms that are repeated three times and over within one minute as the indicator of a chattering alarm (Kondaveeti *et al.* 2013). The alarm chattering may be caused by a single error or the process value that operates close to the alarm threshold. Therefore, the alarm management ensures this kind of alarm should be eliminated or minimized to make the mode of operation viable.

To detect the occurrence of the chattering alarm within the alarm data, we adopted the algorithm from Kondaveeti et al. (Kondaveeti *et al.* 2013) with its run length distribution definition. Run length is the time difference between two consecutive alarms with the same alarm tag. For instance, the alarm that is activated and re-activated with an interval of a second has a 1-second run length. The algorithm explained in Table 3 retrieved the run length distribution from the alarm data set and normalized it with P_r and inversed weighting to calculate the chattering index ψ. The definition of three alarms within one minute is used to defined the chattering index threshold ψ_{cutoff} with value set to 0.05 alarms/second. The alarm tag with chattering index beyond this value is categorized as chattering alarm.

Table 2 Algorithm for chattering index

Algorithm 1: Chattering index

```
Data: Alarm data
Result: chattering index ψ for each alarm tag
ψ_cutoff = 0.05 alarms/second
for alarm tag in alarm data
    r ← run length
    for r in range (1 second, 600 seconds)
        n_r ← counts of alarms with run length r
    end for
    P_r ← n_r / Σ_{r∈N} n_r
    ψ ← Σ_{r∈N} P_r / r
    if ψ > ψ_cutoff
        alarm tag ∈ chattering alarm
    end if
end for
```


https://doi.org/10.24868/10712

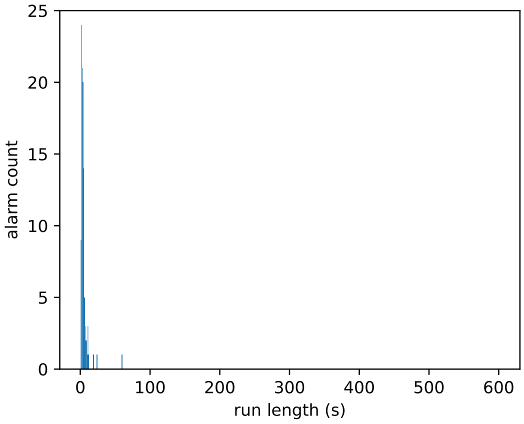

Figure 1 Run length distribution for TAG2505.ABNORMAL

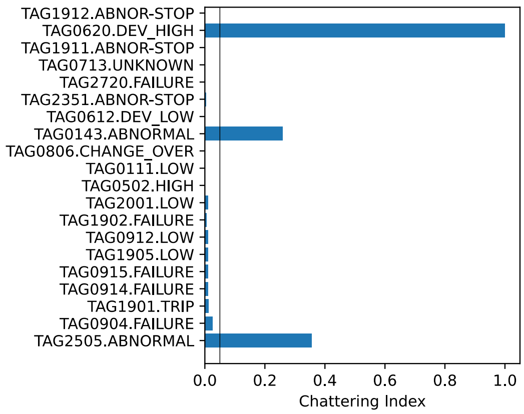

Figure 2 Chattering index for top 20 alarms

Alarm data from the three month voyage length is used as the input. The chattering index was calculated for all unique alarm tags in the alarm data set. Figure 1 shows, for instance, the run length distribution for the alarm with TAG2505.ABNORMAL. The distribution is skewed to the left position; it indicates that several short run lengths have a high number of counts. Figure 2 shows the top 20 alarm tags (by counts on the data set) with their chattering index on the horizontal axis. We indicated three alarms, TAG0620.DEV_HIGH, TAG0143.ABNORMAL and TAG2505.ABNORMAL as the chattering alarm since its chattering index is over the threshold of 0.05 alarms/second.

2.2. Alarm Similarity

Two or more alarms are often activated simultaneously or in random order with a short interval. These kinds of alarms may relate to each other since they may be triggered by the same root cause or the result of inefficient alarm prioritization. Having similar alarms is considered redundant; this condition should be avoided because it increases alarm numbers in the troubleshooting stage. Therefore, the initial step to identify the similarity between the alarm is considered in this study.

We adopted the Jaccard score as an alarm similarity index since it is convenient to be applied in the binary sequence over time (Kondaveeti *et al.* 2012). The alarm similarity index is defined as the measured proximity of occurrence in the time domain when both alarm tags are activated. As shown in Table 3, the first step to calculate the alarm similarity index is to transform the alarm data into a binary sequence for each alarm tag, then padded it with extra binary to cover the communication and time delay on the control system. We adopted the previous study's best practice for using 5 second padded binary before and after the alarm was activated. The

Table 3 Algorithm for alarm similarity index

Algorithm 2: Alarm similarity index

Data: Alarm data
Result: alarm similarity index
for alarm tag in alarm data
    convert into binary
    padded with 5 seconds before and after
end for
for X, Y in alarmtag
    $x_1, x_2, \dots x_N \leftarrow$ alarm X binary sequence
    $y_1, y_2, \dots y_N \leftarrow$ alarm Y binary sequence
    for l in range $(-240, 240)$
        $\mathrm{a}(l) \leftarrow$ count if $x_i = 1$ AND $y_{i+l} = 1$
        $\mathrm{b}(l) \leftarrow$ count if $x_i = 1$ AND $y_{i+l} = 0$
        $\mathrm{c}(l) \leftarrow$ count if $x_i = 0$ AND $y_{i+l} = 1$
        Jaccard score $\leftarrow \mathrm{a}(l)/(\mathrm{a}(l) + \mathrm{b}(l) + \mathrm{c}(l))$
    end for
    similarity index $\leftarrow$ maxValue of Jaccard score
end for

https://doi.org/10.24868/10712

binary sequence is practicable because 0 in the sequence has no meaning since the alarm was not activated. The procedure only considers the time when the alarm is activated at the same time, not dormant at the same time. To make it convenient to visualize, the similarity index is transformed into the distance index, and agglomerative clustering is used to group several alarms with the possibility of similarity.

The alarm similarity index from the top 50 alarm tags is shown in Figure 3. The darker color means a higher alarm similarity index between the alarm pair. Further, clustering can be done here by specifying the threshold. From finding, some clusters consist of alarms with a close alarm tag. For instance: between TAG1908.HIGH, TAG1903.ABNORMAL, TAG1908.LOW; between TAG1413.SENSOR_LOW, TAG1414.HIGH, TAG1414.TROUBLE. It indicates that single trouble may trigger several alarms, and this pattern has recurred several times. The alarm similarity index colormap also indicated the same alarm id with a different tag: for instance, between TAG0616.DEV_LOW and TAG0616.DEV_LOW_S/D. It indicates the alarm is developed into several stages of alarm threshold in close time with a similar pattern, which means the alarm threshold is narrow and may need an adjustment.

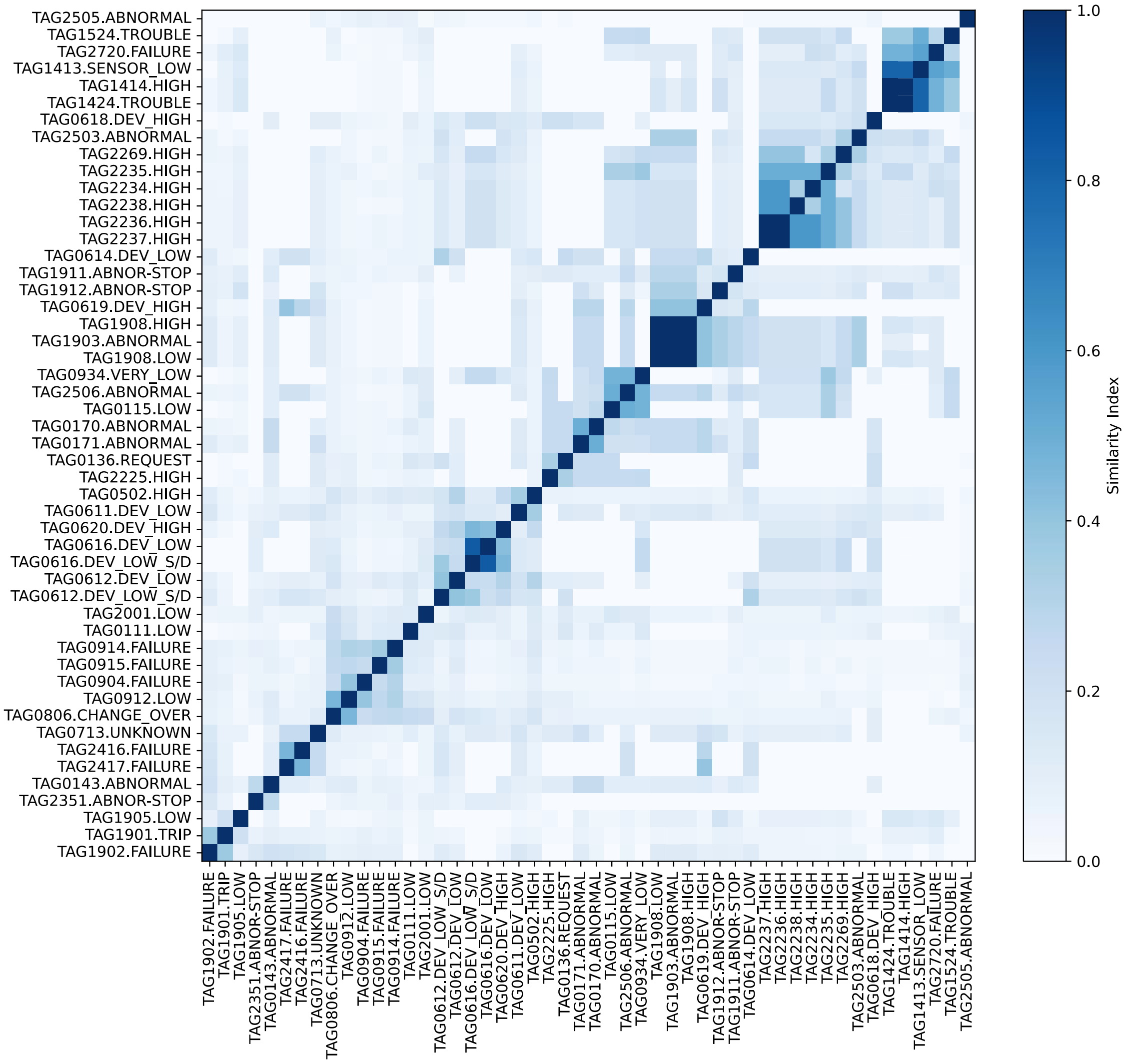

Figure 3 Alarm similarity index colormap

2.3. Alarm Floods Similarity

Multiple alarms may arise during a short period, and troubleshooting condition is possible to be developed into an alarm flood. The guideline of onshore process control defines the alarm flood as the condition where there are ten alarms or more during ten minutes of operation (EEMUA 2009, ANSI/ISA 2016). In this state, the operators cannot handle the alarm effectively because of limited time. Taking the example of the three month voyage length from the same vessel, there are 13 alarm floods after plotting it into 10 minutes bin. The similarity

 https://doi.org/10.24868/10712

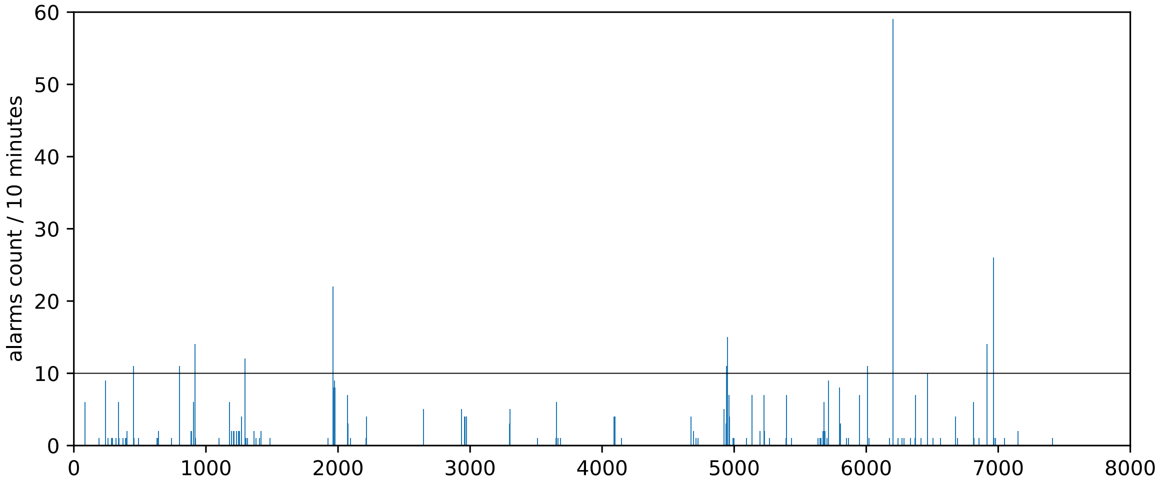

Figure 4 Alarm plot every 10 minutes interval

between these alarm floods may exist because of the same root problem. By identifying activated time and alarm tags in each alarm flood, the similarity can be observed.

We applied the modified Smith-Waterman algorithm, as explained by Cheng et al. (Cheng *et al.* 2013), to conduct local alignment and find the similarity between the alarm floods. The input is the alarm flood sequence identified in the 10-minutes interval plot, as shown in Figure 4. From the alarm flood sequence we retrieved, the alarm in each flood consists of e_m as the unique alarm tag and t_m as its activation time. The next stage calculates the time distance-vector d_m as the distance between mth alarm and the closest alarm on the time domain within the same alarm type k. Weighting vector w_m is applied to discriminate between the closest alarm, first in the sequence, and the alarm that does not exist in the sequence. The similarity index can be calculated for each combination of alarm flood sequences. The sequence pattern between alarm floods can also be retrieved by retracing the matrix H.

Table 4 Alarm Algorithm for alarm flood similarity index

Algorithm 3: Alarm Floods Similarity Index (Modified Smith-Waterman, as in (Cheng *et al.* 2013))

Data: alarm flood sequence database
Result: alarm flood similarity
for alarm flood sequence in alarm flood database
  $X \leftarrow x_1, x_2 \dots x_N$ with N long
  $x_m \leftarrow (e_m, t_m)$, where $m = 1, 2, \dots, M$
  // measuring the distance
  $d_m \leftarrow [d_m^1, d_m^2 \dots d_m^K]^T, for\ k = 1, 2, \dots, K$
  $d_m^k \leftarrow \begin{cases} \min_{1 \le i \le M}\{|t_m - t_i| : e_1 = k\} \\ \infty \end{cases}$
for alarm flood sequence X, alarm flood sequence Y in database
  // weighting
  $w_m \leftarrow [f(d_m^1), f(d_m^2) \dots f(d_m^K)]$
  where, $f(x) \leftarrow e^{-x^2/2\sigma^2}$, for alarm flood sequence X
  $f(x) \leftarrow \begin{cases} 1, & if\ x = 0 \\ 0, & if\ x \neq 0 \end{cases}$, for alarm flood sequence Y
  // calculate the similarity index for each part
  build matrix $H \in R^{(M+1)\times(N+1)}$
  for x in alarm flood sequence X, y in alarm flood sequence Y
    $s\left((e_x, t_x), (e_y, t_y)\right) \leftarrow \max_{1 \le k \le K} |w_x^k \times w_y^k|(1-\mu) + \mu$
  end for
  $H_{p+1,q+1} \leftarrow \max\left(H_{p,q} + s(a_p, b_q), H_{p,q+1} + \delta, H_{p+1,q} + \delta, 0\right)$
  alarm flood similarity $\leftarrow$ maxValue in matrix
end for

https://doi.org/10.24868/10712

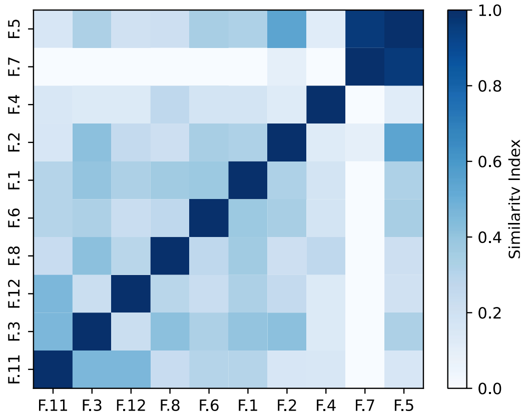

Figure 5 Alarm flood similarity colormap

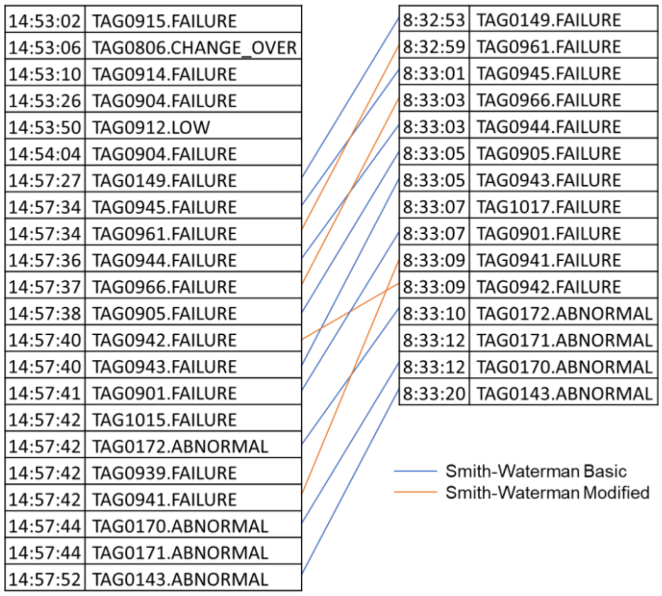

Figure 6 Correlation between Alarm flood

The application of the Smith-Waterman algorithm to analyze the 13 alarm floods is shown in Figure 5. The alarm floods made by single alarm chattering were removed from the analysis. Agglomerative clustering is utilized for clustering. The alarm flood F.5 and F.7, for instance, have high similarity with an index of 0.96. The analysis for the alarm flood sequence gives the extended analysis in Figure 6. The blue line is the result of the basic Smith-waterman algorithm, and the orange line is from the above-modified algorithm. It proved that the modified algorithm by swapping alarm order has more advantages in discovering the similarity between alarm floods.

3. Discussion

Alarm performance assessment was demonstrated as the entry point of the alarm management life cycle. This study applied the guideline from the onshore industry to tackle the absence of comprehensive guidelines in maritime operation. Three measurement indicators were developed based on recent common practice: chattering index, alarm similarity index, and alarm flood similarity index. The actual alarm data from an ocean-going vessel with a three month voyage length was analyzed as an example to visualize the result of each measurement index.

A chattering alarm is considered a nuisance alarm since its presence is unnecessary. More often, it became a burden for the operator and reduced the informativeness of the alarm system. In this analysis, we adopted the definition of run length distribution, the difference between two consecutive alarms with the same alarm tag, to calculate the chattering index for each alarm tag. The unique alarm tag with a chattering index over the defined threshold is then categorized as a chattering alarm. From the data that we retrieved, three alarm tags are indicated as chattering alarms, as shown in Figure 2. Often the chattering alarms in the system arise because the alarm set-points are sensitive. The best practice in alarm management, for instance, is to apply a deadband on the alarm threshold; or if a deadband has been applied, it is to increase its value. To apply it, one must analyze the related sensor's process data to decide the proper deadband value. Once the setting has been applied, we can see the difference before and after by comparing the chattering index. The successful reduction should indicate on lower chattering index, accompanied by the unskewed run length distribution.

The alarm similarity index produced by the Jaccard score, followed by agglomerative clustering, shows the ability to analyze between two or more alarms with similar occurrence time. From the data that we retrieved, several clusters can be defined by examining the heatmap in Figure 3. The possibility that one trouble event may trigger several alarms because the pattern of the occurrence and order has been repeated during the time. Depending on the number of alarms annunciated on a specific vessel, an alarm management practice on the stage of identification and rationalization can be conducted to reduce the redundancy between similar alarms. Another pattern examined from the alarm similarity is that one trouble event may trigger one alarm with a different threshold. It is also advisable to reconfigure the threshold setting.

Thirteen alarm floods condition during the three month voyage length are the input for the alarm floods analysis. The modified Smith-Waterman algorithm shows a more sensitive result by swapping alarm order to

https://doi.org/10.24868/10712

discover the higher similarity between alarm floods. Although the number of alarm floods and clustering results in this study case did not demonstrate any significance, the extended analysis may be considered. For instance, the engineer onboard or onshore staff can examine the root cause to reduce the possibility of the same pattern appearing on the next voyage. Reducing the occurrence of alarm numbers is essential in eliminating the excess information for the crew to handle. When such a pattern is unavoidable, one can also prepare a suitable response strategy.

Three measurement indexes in this study emphasized that the alarm data alone is useful in evaluating the human-machine interface. However, there is a limitation in this study to cover alarm management in more extensive scope. When the process data and process knowledge from the operators is available, it allows the demonstration of alarm optimization and rationalization technique such as filtering, deadbands, and delay times. These techniques are practicable for existing ship control systems. However, with similar machinery and dimension, a newly built sister ship can adopt the alarm management practice from the previously constructed ship. The absence of specific guidelines in maritime operations urges that comprehensive guidelines, such onshore industry has been done, should also be introduced in maritime operations. The alarm performance or alarm management is beneficial not only for the current operation but also for future projection of remote and autonomous ship where alarm indicator plays essential roles.

Acknowledgments

This study was supported by JSPS KAKENHI Grant No. 21K04517.

References

Cheng, Y., Izadi, I., and Chen, T., 2013. Pattern matching of alarm flood sequences by a modified Smith–Waterman algorithm. *Chemical Engineering Research and Design*, 91 (6), 1085–1094.

Chilcott, J. and Kennedy, N., 2018. Enabling Lean Manning through Automation. *In*: *International Ship Control Systems Symposium*. Presented at the International Ship Control Systems Symposium, Glasgow, UK.

Danish Maritime Accident Investigation Board, 2013. *Marine accident report: Emma Maersk flooding of engine room on 1 February 2013*. Denmark, No. 2013002165.

Engineering Equipment and Materials Users Association, 2009. *EEMUA Publication 191 Alarm systems - a guide to design, management and procurement*.

International Maritime Organization, 2009. *Code on Alerts And Indicators*. No. A 26/Res.1021.

International Maritime Organization, 2010. *Adoption Of Performance Standards For Bridge Alert Management*. No. MSC 87/26/Add.1.

International Society of Automation and American National Standards Institute, 2016. *Management of alarm systems for the process industries*. ANSI/ISA, No. ANSI/ISA–18.2–2016.

Jones, B.S., Earthy, J.V., Fort, Ed., and Gould, D., 2006. Improving the design and management of alarm systems. *In*: *World Maritime Technology Conference*. 2006: IMarEST, 7.

Kondaveeti, S.R., Izadi, I., Shah, S.L., Black, T., and Chen, T., 2012. Graphical tools for routine assessment of industrial alarm systems. *Computers & Chemical Engineering*, 46, 39–47.

Kondaveeti, S.R., Izadi, I., Shah, S.L., Shook, D.S., Kadali, R., and Chen, T., 2013. Quantification of alarm chatter based on run length distributions. *Chemical Engineering Research and Design*, 91 (12), 2550–2558.

Lundh, M., Lützhöft, M., Rydstedt, L., and Dahlman, J., 2011. Working conditions in the engine department – A qualitative study among engine room personnel on board Swedish merchant ships. *Applied Ergonomics*, 42 (2), 384–390.

Man, Y., Lundh, M., and MacKinnon, S.N., 2018. Managing unruly technologies in the engine control room: from problem patching to an architectural thinking and standardization. *WMU Journal of Maritime Affairs*, 17 (4), 497–519.

Rødseth, Ø.J., Knight, M., Storari, R., Foss, H., and Tinderholt, A.R., 2006. Alarm management on merchant ships. *In*: *World Maritime Technology Conference*. London, 10.