OPERATION ANALYSIS OF ETHYLENE PLANT BY EVENT CORRELATION ANALYSIS OF OPERATION LOG DATA

Masaru Noda*[1], Tsutomu Takai[1] and Fumitaka Higuchi[2]
[1] Nara Institute of Science and Technology,
Nara 630-0192, Japan
[2] Idemitsu Kosan Co. Ltd.,
Chiba 261-8501, Japan

Abstract

Event correlation analysis is a method of extracting knowledge that detects statistical similarities between discrete events. The method can identify unnecessary alarms and operations from the operation log data of chemical plant. In the improved method of the event correlation analysis, the time window is expanded, and the log data of two events are reconverted into sequential binary data using the updated size of the time window, when a high degree of similarity between two events is not detected. The time window continues to be expanded and similarity continues to be recalculated until either a high degree of similarity is detected or the time window becomes larger than the maximum pre-determined size. We applied the improved event correlation analysis to the operation data of an ethylene plant. The results revealed that it was able to correctly identify similarities between two physically related events, even when the conventional method using a constant time-window size failed due to the large variance in time lag. Unnecessary alarms and operations within a large amount of event data from industrial chemical plants could effectively be identified using the new method.

Keywords

Alarm management, plant alarm system, event correlation analysis, operation log data, ethylene plant

Introduction

The progress with distributed control systems (DCSs) in the chemical industry has made it possible to install many alarms easily and inexpensively. While most alarms help operators detect abnormalities and identify their causes, some are unnecessary. A poor alarm system might cause floods of alarms and nuisance alarms, which would reduce the ability of operators to cope with abnormalities at plants because critical alarms were buried under many that were unnecessary (Nimmo, 2002, Alford, 2005).

The independent protection layers (AIChE/CCPS, 1993) summarized in Table 1 have been extensively applied to various types of plants to protect them from hazardous incidents. Alarm systems, which are located at the third layer of the independent protection layers, activate alarms to notify operators to take corrective action when the process deviates from normal operating conditions.

The Engineering Equipment and Materials Users' Association (EEMUA, 2007) defined the primary function of an alarm system as directing the operator's attention toward plant conditions requiring timely assessment or action. To achieve this, every alarm should have a defined response and adequate time should be allowed for the operator to carry out this response. The International

*Corresponding author's e-mail: noda@is.naist.jp

Society of Automation (ISA, 2009) suggested a standard alarm-management lifecycle covering alarm-system specifications, design, implementation, operation, monitoring, maintenance, and change activities from initial conception through to decommissioning. The lifecycle model recommends the continuous monitoring and assessment of operation log data to rationalize alarm systems.

Table 1 Independent protection layers for process safety (AIChE/CCPS, 1993)

Layers	Definitions
8	Community emergency response
7	Plant emergency response
6	Physical protection (Dikes)
5	Physical protection (Relief devices)
4	Automatic action SIS or ESD
3	Critical alarms, operator supervision, and manual intervention
2	Basic controls, process alarms, and operator supervision
1	Process design

The "top-ten worst alarm method" has been widely used in the chemical industry to reduce the number of unnecessary alarms. It is used to collect data from the event logs of alarms during operation and it creates a list of frequently generated alarms. The alarms are then reviewed one after another, starting with the one most frequently triggered, and the root causes that triggered them are identified. Although this method can effectively reduce the number of alarms triggered at an early stage, it is less effective at reducing them as the proportion of the worst ten alarms decreases. Because the ratio of each alarm in the top-ten worst alarm list is very small in the latter case, it becomes difficult to achieve further effective improvements.

Kondaveeti *et al.* (2009) proposed the High Density Alarm Plot (HDAP) and the Alarm Similarity Color Map (ASCM) to assess the performance of alarm systems in terms of effectively reducing the number of nuisance alarms. HDAP visualizes the time various alarms occurred, which facilitated the identification of periods when the plant was unstable. ASCM orders alarms according to their degree of Jaccard similarity (Lesot *et al.*, 2009) with other alarms to identify redundant alarms. However, these visualization tools are not able to designate whether individual alarms have a defined response, because they only focus on alarms in the operation log data.

Nishiguchi and Takai (2010) proposed a method of data-based evaluation that referred to not only alarm event data but also operation event data in the operation log data of plants. It used event correlation analysis to detect statistical similarities between discrete alarms or operation events. Grouping correlated events based on their degree of similarity made it possible to consider countermeasures to reduce the frequency of alarms more easily than could be done merely by analyzing individual alarms and operation events. Event correlation analysis was applied to the operation log data of industrial chemical plants. Unnecessary alarms and operations were accurately identified within a large amount of event log data by using the method (Higuchi *et al.*, 2010). However, it occasionally failed to detect similarities between two physically related events when there was too much variance in the time lag between them.

Kurata *et al.* (2011) improved event correlation analysis, which was able to detect similarities between physically related events with large variance in time lag. The time window in their method was expanded, and the log data of two events were reconverted into sequential binary data using the new time-window size when high degrees of similarity between two events were not detected. The time window continued to be expanded and similarity continued to be recalculated until either a high degree of similarity was detected or the time window became larger than the maximum pre-determined size.

We applied the improved method of event correlation analysis to the operation log data of an ethylene plant operated by Idemitsu Kosan Co. Ltd. in Japan to test and confirm whether the method was able to correctly identify similarities between two physically related events.

Improved Event Correlation Analysis (Kurata *et al.*, 2011)

The plant log data recorded in DCS consist of the times of occurrences and the tag names of alarms or operations as listed in Table 2, which we will call "events" after this.

Table 2 Example of event log data

Date/Time	Event	Type
2011/01/01 00:08:53	Event 1	Alarm
2011/01/01 00:09:36	Event 2	Operation
2011/01/01 00:11:42	Event 3	Alarm
2011/01/01 00:25:52	Event 1	Alarm
2011/01/01 00:30:34	Event 2	Operation
:	:	

First, the plant log data are converted into sequential event data $s_i(k)$ by using Eq. (1). When event i occurs between $(k\text{-}1)\Delta t$ and $k\Delta$t, $s_i(k) = 1$, otherwise $s_i(k) = 0$. Here, Δt is the time-window size and k denotes the discrete time. Figure 1 has an example of a binary sequence of event log data.

$$s_i(k) = \begin{cases} 1 \text{ if event } i \text{ occurs between } (k-1)\Delta t \text{ and } k\Delta t \\ 0 \text{ otherwise} \end{cases} \quad (1)$$

$$(1 \le k \le T/\Delta t)$$

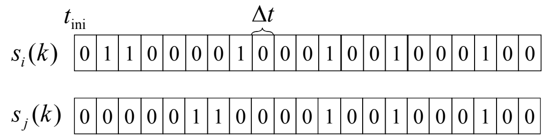

Fig. 1 Binary sequence of event log data

The cross correlation function, $c_{ij}(m)$, between $s_i(k)$ and $s_j(k)$ for time lag m is calculated with Eq. (2). Here, K is the maximum time period for lag and T is the time period for whole event log data.

$$c_{ij}(m) = \begin{cases} \sum_{k=1}^{T/\Delta t - m} s_i(k) s_j(k+m) & (0 \le m \le K/\Delta t) \\ c_{ji}(-m) & (-K/\Delta t \le m < 0) \end{cases} \tag{2}$$

The maximum value of cross correlation function c_{ij}^* is obtained with Eq. (3).

$$c_{ij}^*(m) = \max_m c_{ij}(m) \tag{3}$$

Here, we assumed that two events i and j are independent of each other. If probability p_{ij} that two events i and j will occur simultaneously is very small, the probability distribution that two events will occur simultaneously is approximated by the Poisson distribution. The total probability that two events will occur simultaneously more than c_{ij}^* times with time lag m is given by Eq. (4), where λ is the expected value of c_{ij} (Mannila and Rusakov, 2001).

$$P(c_{ij}(m) \ge c_{ij}^* | -K/\Delta t \le m \le K/\Delta t) \cong 1 - \left(\sum_{l=0}^{c_{ij}^* - 1} \frac{e^{-\lambda} \lambda^l}{l!} \right)^{2K+1} \tag{4}$$

Finally, the similarity, S_{ij}, between two events i and j is calculated with Eq. (5) (Nishiguchi and Takai, 2010).

$$S_{ij} = 1 - P(c_{ij}(m) \ge c_{ij}^* | -K/\Delta t \le m \le K/\Delta t) \tag{5}$$

If a high degree of similarity between two events is not detected, the time window is doubled in size by using Eq. (6), and the log data of two events are reconverted into sequential binary data using the new time-window size, as seen in Fig. 2 (Kurata *et al.*, in press). The time window continues to be expanded and similarity continues to be recalculated until either a high degree of similarity is detected or the time window becomes larger than the maximum pre-determined size, $\Delta t_{\max}$.

$$s_i'(k) = \begin{cases} 1 & \text{if } s_i(2k-1) = 1 \vee s_i(2k) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$(1 \le k \le T/\Delta t')$$

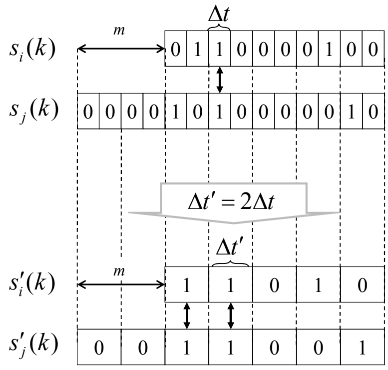

Fig. 2 Updating time window size (Kurata *et al.*, 2011)

A larger similarity means a stronger dependence or closer relationship between the two events. After similarities are calculated between all combinations of any two events in the plant log data, all events are classified into groups with a hierarchical method of clustering, where the distance between two events i and j is defined by Eq. (7). It becomes possible to stratify and visualize the distance between events by grouping them.

$$D_{ij} = 1 - S_{ij} \tag{7}$$

The following four types of nuisance alarms and operations can be found by analyzing the results obtained from clustering.

(1) **Sequential alarms**: When a group contains multiple alarm events that occur sequentially, these are sequential alarms. Changing the alarm settings of sequential alarms may effectively reduce the number of times they occur.
(2) **Routine operations**: When many operation events are included in a group and operation events in the same group appear frequently in the event log data, they may be routine operations. These operation events can be reduced by automating routine operations using a programmable logic controller.
(3) **Alarms without corresponding operations**: When there are only alarm events in a group and operation events are not included in the same group, they may be alarms without corresponding operations. As every alarm should have a defined response (EEMUA, 2009), these may be unnecessary and should be eliminated.
(4) **Alarms after operation**: Alarm events occur after all operation events in a group, and these may be caused by operations. These are unnecessary because they are not meaningful or actionable.

Operation Log Data of Ethylene Plant

Idemitsu Kosan Co. Ltd. started operations at the ethylene plant of their Chiba complex in 1985. Figure 3 is a process flow diagram for the ethylene plant, which is operated by two board operators using DCS. The plant IDs in Fig. 3 indicate the identification number of plants, which are summarized in Table 3.

The total numbers of alarm events and operation events in DCS correspond to 3236 and 775 for process control and process monitoring. When an alarm or operation event occurs, the event name and the occurrence time are recorded in the operation log data every minute in DCS.

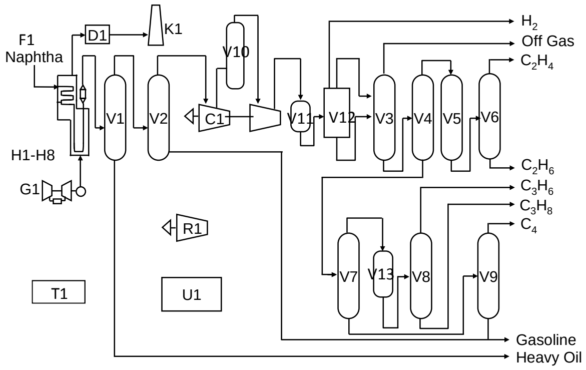

Fig. 3 Process flow diagram for ethylene plant (Higuchi *et al.*, 2010)

Table 3 Units in ethylene plant

No.	Unit name	No.	Unit name
C1	Cracked gas compressor	V2	Quench water tower
D1	DeNOx section	V3	Demethanizer
F1	Feed	V4	Deethanizer
G1	Gas turbine	V5	Acetylene absorber
H1–H8	Cracking furnaces 1–8	V6	Ethylene fractionator
K1	Exhaust gas stack	V7	Depropanizer
P1	Product processing unit	V8	Propylene fractionator
R1	Refrigeration compressor	V9	Debutanizer
T1	Tank	V11	Dryer
U1	Utility section	V12	Chill train
V1	Primary fractionator	V13	Hydrogenation Reactor

The plant log data gathered in one month included 914 types of alarm events and 857 types of operation events. A total number of 51640 events was generated. Figure 4 shows the points at which 1771 types of alarm and operation events occurred. It is difficult to identify sequential alarms, and alarms without corresponding operations, by merely scrutinizing Fig. 4. Figure 5 shows the frequency of alarm events generated in the ethylene plant over ten minutes. Idemitsu Kosan Co., Ltd. applied the top-ten worst alarm method to the problem to decrease alarm rates as part of its total maintenance activities during production. However, the ethylene plant could not in fact achieve EEMUA's guidelines of an average-alarm-frequency standard during normal operations.

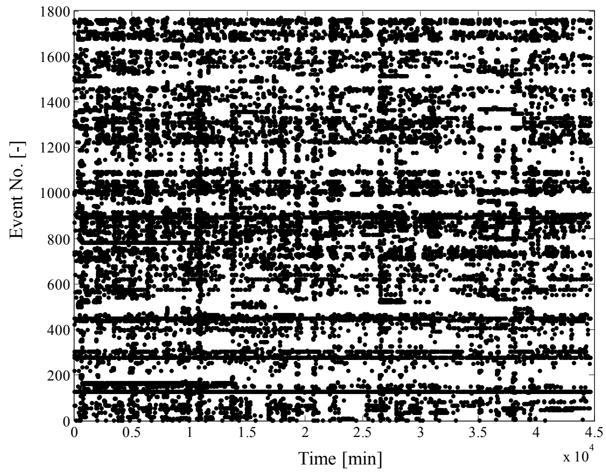

Fig. 4 Event log data for ethylene plant

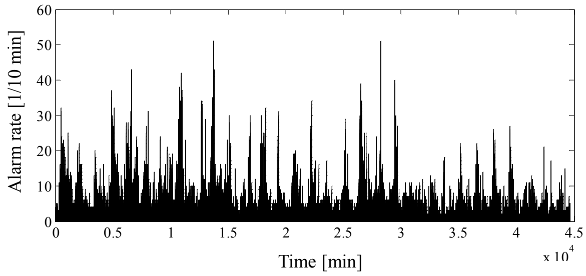

Fig. 5 Frequency of alarms generated in ethylene plant

Results from Event Correlation Analysis

Event correlation analysis was applied to the operation log data obtained from the ethylene plant, where the minimum threshold to identify similarities between two events was set at 0.995. By using the hierarchical method of clustering, 1771 types of alarms and operation events were classified into 588 groups. The worst 10 groups are summarized in Table 4. Figure 6 is an alarm similarity color map of events in the top 10 worst groups, where the alarm and operation events are ordered according to the group Nos. The red in Fig. 6 indicates that two events have a high degree of similarity between them. The alarm similarity color map is extremely helpful for identifying related alarms and operations at a glance.

The top group contains five types of alarm events and ten types of operation events, and the total number of events in the group accounted for 5.8% of all generated events at the ethylene plant. Although the total number of events in the worst 10 groups accounted for 32.4% of all generated events at the plant, only 4.2% of alarm and operation event types were in them.

Table 4 Top 10 worst groups

Group No.	Number of events			Number of types	
	Total	Alarm	Operation	Alarm	Operation
1	2983	212	2771	5	10
2	2377	2377	0	2	0
3	1795	938	857	1	2
4	1693	25	1668	1	6
5	1585	1585	0	2	0
6	1507	241	1266	4	7
7	1290	0	1290	0	8
8	1243	0	1243	0	6
9	1214	32	1182	2	8
10	1049	118	931	4	6

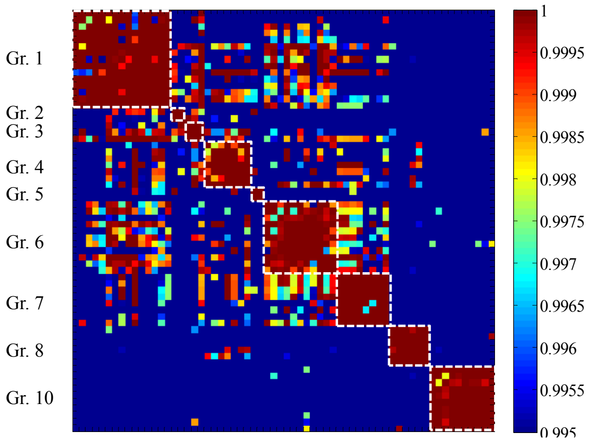

Fig. 6 Alarm similarity color map for top 10 worst groups

Groups 2 and 5 only contained alarm events, which means that these alarm events were not followed by corresponding operations. According to EEMUA's key design principles for alarm systems, every alarm should have a defined response. Sometimes the response to the alarm is conditional, e.g., an operator may only carry out a defined response in certain circumstances. If a response cannot be defined for alarm events in groups 2 and 5, these alarms should be removed.

Groups 7 and 8 only consisted of operation events and these operation events occurred more than thousand times in one month. When many operations are included in a group, these may be routine operations. Routine operations can be eliminated by implementing an intelligent system to control sequences.

Except for groups 3, 4, 7, and 8, all groups contained multiple alarms. These alarms were supposed to be sequential alarms. Sequential alarms distract operators by raising multiple alarms caused by single events. Only one such alarm should be configured at the point where the operator is most likely to take action (Hollifield and Habibi, 2006).

Changing the alarm settings according to the state of the plant, improving the performance of controls, and automating operations by using sequence-control programs reduced number of alarms and operations included in the worst 10 groups. Implementing a programmable logic controller, in which alarm settings were automatically changed according to the state of the plant and operations, significantly decreased the large number of events generated by operations in an unsteady state.

Conclusion

The improved method of event correlation analysis was applied to the plant operation data of an ethylene plant. The results demonstrated that it was able to correctly identify similarities between two physically related events, even when the conventional method using a constant time window size failed due to the large variance in time lag. We could effectively identify unnecessary alarms and operations within a large amount of event data by using the method, which would be helpful for reducing the number of unnecessary alarms and operations in other industrial chemical plants.

Acknowledgments

The authors gratefully acknowledge the cooperation extended by Idemitsu Kosan Co. Ltd. in providing us with invaluable data from their ethylene plant.

References

AIChE/CCPS, (1993). Guidelines for Engineering Design for Process Safety. AIChE, New York, NY.

Alford, J. S., Kindervater, K., Stankovich, R., (2005). Alarm Management for Regulated Industries, *Chemical Engineering Progress*, 101, 25.

Engineering Equipment & Material Users' Association (2007). Alarm Systems - A Guide to Design, Management and Procurement, EEMUA Publication No.191 2nd Edition, EEMUA, London

Higuchi, F., Noda, M., Nishitani H. (2010). Alarm Reduction of Ethylene Plant using Event Correlation Analysis (in Japanese), *Kagaku Kogaku Ronbunshu*, 36, 576

Hollifield, B. R., Habibi, E. (2009). The Alarm Management Handbook. PAS. Houston, TX.

Kurata, K., Noda, M., Kikuchi, Y. Hirao, M. (2011). Extension of Event Correlation Analysis for Rationalization of Plant Alarm Systems (in Japanese), *Kagaku Kogaku Ronbunshu*, 37, 338.

Kondaveeti, S. R., Shah S. L., Izadi, I. (2009). Application of Multivariate Statistics for Efficient Alarm Generation, *Proc. of 7th IFAC Symposium of Fault Detection, Supervision and Safety of Technical Processes*, 657.

Lesot, M. J., Rifqi M., Benhadda H. (2009). Similarity measures for binary and numerical data: a survey, *Int. J. Knowledge Engineering and Soft Data Paradigms*, 1, 63.

Mannila, H., Rusakov, D. (2001). Decomposition of Event Sequences into Independent Components, *Proc. of 2001 SIAM International Conferences on Data Mining*.

Nimmo, I. (2002). Consider Human Factors in Alarm Management, *Chemical Engineering Progress*, 98, 30.

Nishiguchi, J., Takai, T. (2010). IPL2 and 3 performance improvement method for process safety using event correlation analysis. *Computers & Chemical Engineering*, 34, 2007.