An Overview of Industrial Alarm Systems: Main Causes for Alarm Overloading, Research Status, and Open Problems

Jiandong Wang, Fan Yang, Tongwen Chen, and Sirish L. Shah

***Abstract*—Alarm systems play critically important roles for the safe and efficient operation of modern industrial plants. However, most existing industrial alarm systems suffer from poor performance, noticeably having too many alarms to be handled by operators in control rooms. Such alarm overloading is extremely detrimental to the important role played by alarm systems. This paper provides an overview of industrial alarm systems. Four main causes are identified as the culprits for alarm overloading, namely, chattering alarms due to noise and disturbance, alarm variables incorrectly configured, alarm design isolated from related variables, and abnormality propagation owing to physical connections. Industrial examples from a large-scale thermal power plant are provided as supportive evidences. The current research status for industrial alarm systems is summarized by focusing on existing studies related to these main causes. Eight fundamental research problems to be solved are formulated for the complete lifecycle of alarm variables including alarm configuration, alarm design, and alarm removal.**

***Note to Practitioners*—Alarm systems are critical assets for operational safety and efficiency of plants in various industrial sectors, such as power and utility, process and manufacturing, and oil and gas. However, industrial alarm systems are generally suffering from alarm overloading. This paper provides an overview of industrial alarm systems, by proposing main causes for alarm overloading, summarizing current research status and formulating open problems. In presenting this overview, we hope to attract direct attentions from more researchers and engineers into the study of industrial alarm systems.**

***Index Terms*—Alarm configuration, alarm design, alarm removal, industrial alarm systems, nuisance alarms.**

Manuscript received January 19, 2015; revised June 05, 2015; accepted July 26, 2015. Date of publication September 16, 2015; date of current version April 05, 2016. This paper was recommended for publication by Associate Editor J.-M. Faure and Editor M. Wang upon evaluation of the reviewers' comments. This work was supported by the National Natural Science Foundation of China under Grant 61433001, and in part by the Natural Sciences and Engineering Research Council of Canada under Grant CRDPJ-446412-12. *(Corresponding author: Jiandong Wang.)*

J. Wang is with the College of Engineering, Peking University, Beijing 100871, China (e-mail: jiandong@pku.edu.cn).

F. Yang is with the Tsinghua National Laboratory for Information Science and Technology and the Department of Automation, Tsinghua University, Beijing 100084, China (e-mail: yangfan@tsinghua.edu.cn).

T. Chen is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: tchen@ualberta.ca).

S. L. Shah is with the Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: sirish.shah@ualberta.ca).

Digital Object Identifier 10.1109/TASE.2015.2464234

I. Introduction

According to the industrial standard ANSI/ISA-18.2 [58, p. 16], "an alarm system is the collection of hardware and software that detects an alarm state, communicates the indication of that state to operators, and records changes in the alarm state." Alarm systems have been the integrated parts of modern computerized monitoring systems such as the distributed control systems (DCS) and supervisory control and data acquisition (SCADA) systems. The most common way in detecting an alarm state is to compare the value of a process variable $x(t)$ to a constant high (low) alarm trippoint x_{tp}, i.e.,

$$x_a(t) = \begin{cases} 1, & \text{if } x(t) \geq (\leq)\, x_{tp} \\ 0, & \text{if } x(t) < (>)\, x_{tp} \end{cases}. \tag{1}$$

As an example, $x(t)$ is the drum level of a large-scale thermal power plant at Shandong Province, China, referred to as Plant A in the sequel, associated with the high alarm trippoint 100 and the low alarm trippoint -100. Fig. 1(a) presents 1-hour samples of $x(t)$ with sampling period 1 s. Alarm states in alarm systems can be mathematically represented by a discrete-valued alarm variable $x_a(t)$. Fig. 1(b) shows the samples of two alarm variables associated with $x(t)$ in Fig. 1(a). That is, the high (low) alarm variable $x_a(t)$ takes the value of 1 when $x(t)$ is higher (lower) than the value 100 (-100), and the value of 0, otherwise. The changes of alarm variables from 0 to 1 and from 1 to 0 are respectively referred to as the alarm occurrence and alarm clearance. They appear as events in the alarm list with the corresponding time stamps and descriptions, as shown in Table I. Note that the low alarm variable experienced two quick changes between 23:47:38 and 23:47:42, which are visible by the enlarged plot in Fig. 1(b).

Alarm systems are critically important for safe and efficient operations of modern industrial plants such as oil refineries, petrochemical facilities, and power plants [17], [89]. First, alarm systems are the tools to detect the near misses that are defined as departures from and subsequent returns to normal operating ranges for process variables [88]. For instance, the variations of two alarm variables in Fig. 1 indicate that the drum level departs from and returns to the normal operating range $[-100, 100]$. The safety pyramid in Fig. 2 says that every accident is associated with a number of near misses as precursors. Alarm systems promptly indicate the occurrences of near misses, so that operators can take corrective actions to drive

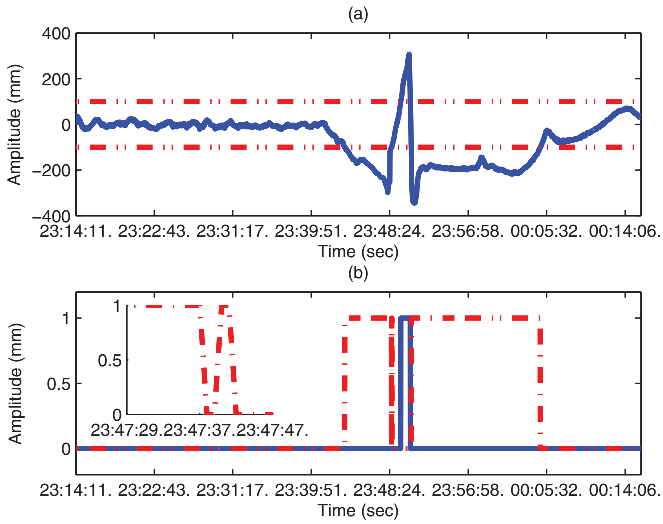


Fig. 1. (a) Collected samples of a process variable $x(t)$ (solid) with alarm trippoints (dot-dash). (b) Alarm variables $x_a(t)$ with high (solid) and low (dot-dash) alarm trippoints.

TABLE I
EVENTS OF ALARM OCCURRENCES AND CLEARANCES FOR HIGH AND LOW DRUM LEVEL ALARM VARIABLES IN FIG. 1

Date	Time	Status	Tagname	Description
2014-01-24	23:42:42	Alarm on	DRUMLVLL	Drum level low
2014-01-24	23:47:38	Alarm off	DRUMLVLL	Drum level low
2014-01-24	23:47:40	Alarm on	DRUMLVLL	Drum level low
2014-01-24	23:47:42	Alarm off	DRUMLVLL	Drum level low
2014-01-24	23:48:38	Alarm on	DRUMLVLH	Drum level high
2014-01-24	23:49:39	Alarm off	DRUMLVLH	Drum level high
2014-01-24	23:49:45	Alarm on	DRUMLVLL	Drum level low
2014-01-25	00:03:27	Alarm off	DRUMLVLL	Drum level low

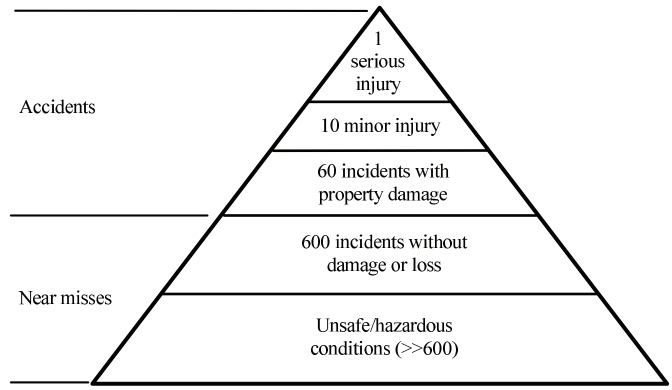


Fig. 2. Safety pyramid with typical historical data [88].

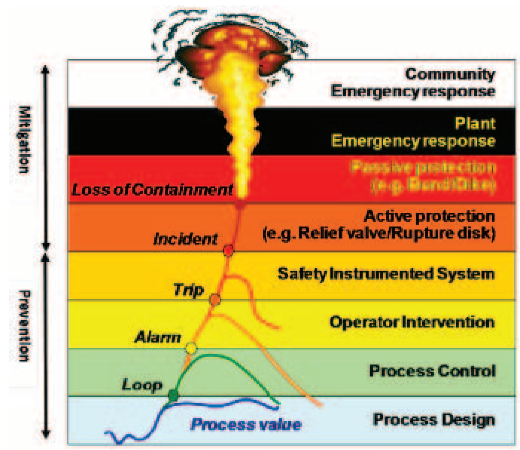


Fig. 3. Layers of protection and their impact [94].

processes back to normal operating ranges. Therefore, alarm systems are indeed the safeguards to prevent the deterioration of near misses to accidents. Second, retrospective investigation on a large number of accidents support the important role played by alarm systems. For instance, the Buncefield accident in 2005 was caused by a series of explosions and fire of oil storages at the Buncefield oil depot, and is by far the most severe industrial accident in Europe. The final report of the Buncefield accident [54] provided 25 recommendations, among which the 8th recommendation was to develop high–high level alarms for overfill prevention, and the 23rd recommendation was to collect accident data to find alarm system defects. Finally, alarm systems also play a prominent role in maintaining efficiency of plant operation. It is a well-known fact that the deviation of process variables from normal/optimal operating zones usually imply negative effects such as off-specification products as well as excessive consumption of raw materials and energy. Alarm systems with satisfactory performance are able to assist operators to reduce the probability and time duration of deviations of key process variables from their normal/optimal operating zones. Overall, as shown in Fig. 3, alarm systems play significant roles in maintaining plant safety and keeping the process within normal operating ranges [77], [94].

On one hand, alarm systems are critically important for safety and efficiency of industrial plants; on the other hand, most existing industrial alarm systems suffer from poor performance, most noticeably with alarm overloading (to be clarified later in Section II). Driven by the big gap between these facts, industrial standards and guidelines have been proposed for alarm systems by industrial societies and professional organizations. The Nuclear Regulatory Commission from the United States published the document NUREG/CR-61056684 to give guidance and technical basis for advanced alarm systems [87]. The Engineering Equipment and Materials Users' Association presented the guideline EEMUA-191 for the design, management and procurement of alarm systems [41]. The Standardization Association for Measurement and Control in Chemical Industries issued the standard NAMUR-NA-102 for alarm management [83]. The USA-based Electric Power Research Institute produced the document EPRI-1010076 as the requirements and implementation guidance for advanced alarm systems [42]. The Abnormal Situations Management Consortium proposed a set of guidelines for effective alarm management practice [10]. The International Society of Automation and the International Electrotechnical Commission, respectively, issued the standards ANSI/ISA-18.2 [58] and IEC-68682, for management of alarm systems in process industries. The American Petroleum Institute published the standard API-1167 for pipeline SCADA (supervisory control and data acquisition) alarm management [7]. Note that all of these standards and guidelines impose specific requirements on the performance of alarm systems, for example, "the average number of alarms per day should be no more than 144 [41]," but do not provide methodologies and/or technical details on how to achieve these requirements.

The importance of studying alarm systems for performance improvement has also received approval from completed/ongoing research projects. The European Commission's Community Research and Development Information Service supported the Fifth Framework Programme titled "Advanced decision support system for chemical/petrochemical manufacturing processes (2001–2004)," where some toolboxes have been developed for alarm management. The Australian Research Council supported a project titled "Alarm management: silence is golden (2003–2007)." The Western Australia Energy Research Alliance supported a project titled "Adaptive optimization learning applied to real time alarm management (2008–2012)." Japan Society for the Promotion of Science formulated the 143rd Committee on Process System Engineering that supported a project titled "Alarm management (2009–2011)." The Natural Sciences and Engineering Research Council of Canada supported a strategic grant titled "Development of an advanced technology for alarm analysis and design (2009–2012)." The National Natural Science Foundation of China is supporting a project titled "Alarm design and removal methods and applications for highly efficient and safe operation of large-scale industrial systems (2015–2019)."

This paper provides an overview of industrial alarm systems with the following contributions: (i) four main causes are identified as the culprits for alarm overloading in many industrial alarm systems; industrial examples from Plant A are provided as supportive evidences; (ii) the current research status for industrial alarm systems is summarized, by focusing on the existing studies related to these main causes; (iii) eight fundamental research problems, which are still open, are formulated for the complete lifecycle of alarm variables composed of three stages, namely, alarm configuration, alarm design and alarm removal. In presenting this overview, one of our objectives is to attract direct attentions from more researchers and engineers into the fascinating area of industrial alarm systems.

~~The remaining sections of this paper are organized as follows. Section II investigates the current status of industrial alarm systems. Section III identifies four main causes of alarm overloading. Section IV summarizes the research status of industrial alarm systems. Section V formulates fundamental research problems to be solved. Section VI gives some concluding remarks~~.

II. Current Status of Alarm Systems: Alarm Overloading

In this section, we investigate the current status of alarm systems, based on industrial surveys in literature and the status of Plant A.

Many existing industrial alarm systems are associated with poor performance, where the most observable phenomenon is that there are far too many alarms to be handled by industrial plant operators, referred to as alarm overloading in the sequel. This phenomenon is clearly revealed from Table II [89], which lists statistics of several basic performance metrics of alarms systems, based on a study of 39 industrial plants ranging from oil and gas, petrochemical, power and other industries. The corresponding benchmarks in the guideline EEMUA-191 [41] are also provided in Table II for comparison. Obviously, the statistics of performance metrics from various industries are much greater than the EEMUA benchmarks. Another industrial survey was provided by Bransby and Jenkinson [17] for 15 plants including oil refineries, chemical plants, pharmaceutical plants, gas terminal, and power stations; the average alarm rate per 10 min under normal operation ranged from 2 to 33, and the peak alarm rate per 10 min in plant upsets varied from 72 to 625. Brown [20] provided similar results for BP Oil plants: the average rate of alarms per 10 min was in the range of [17], [60], while the maximum rate of alarms per 10 min in upset conditions was in the range of [150, 560]. Noda *et al.* [86] reported that 15 out of 29 Japanese chemical plants had the monthly average alarm rates per operator larger than the EEMUA benchmark, with the maximum value of 7.5 alarms per operator for 10 min, where the number of operators was taken into consideration. The number of operators was also recorded for the industrial survey by Bransby and Jenkinson [17]; for most of cases, however, the statistics of alarm systems were not normalized by the number of operators. Kirschen and Wollenberg [66] presented the estimates of peak numbers of alarms triggered by some abnormal events at a regional control centers of Hydro Quebec, Inc., Canada, which were up to 20 alarms per second during a thunderstorm. Liu *et al.* [75] and Srinivasan *et al.* [93] stated that the daily alarm number in abnormal operation was about 11000 (76 alarms per 10 min) at a major Singapore refinery.

TABLE II
Cross-Industry Study [89]

	EEMUA	Oil-Gas	PetroChem	Power
Average alarms/day	144	1200	1500	2000
Peak alarms/10 min	10	220	180	350
Average alarms/10 min	1	6	9	8

To have an example of industrial alarm systems with more specifics, we investigated the alarm system of a 300 MW power generation unit at Plant A. The distributed control system (DCS) of the power generation unit measures real-time values of 24079 process variables every 0.2 s. Among these process variables, there are 8145 analog variables among which 158 variables are configured with alarms, and 15934 digital variables among which 1784 variables are configured with alarms. An alarm occurs when an analog variable configured with alarm is exceeding the corresponding high (high–high) or low (low–low) alarm trippoint, or when a digital variable configured with alarm changes the value from 0 to 1 (or from 1 to 0). It is worthy to note that some digital variables configured with alarms are essentially generated by comparing the measurements of analog variables with alarm trippoints. Fig. 4 presents the numbers of alarm occurrences during non-overlapping consecutive 10-min periods for 31 days in March 2014. The

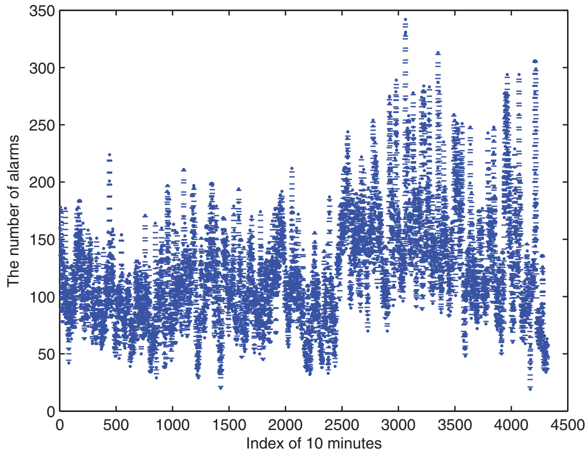


Fig. 4. The number of alarms per 10 minutes of one process unit at Plant A for 31 days in March 2014.

maximum, average and minimum numbers of the alarm occurrence rate per 10 min are 342, 119 and 19, respectively. Clearly, the alarm occurrence rates are much larger than the EEMUA benchmarks in Table II.

The occurred alarms can be classified into two groups, namely, nuisance alarms and correct alarms. A nuisance alarm is one that does not require a specific action or response from operators [41], [89]. An alternative definition of a nuisance alarm is an alarm that annunciates excessively, unnecessarily, or does not return to normal after the correct response is taken [58, p. 18]. Hence, nuisance alarms in a process are defined as the occurred alarms that do not affect the process, even if these alarms are ignored by operators. The industrial standard ANSI/ISA-18.2 [58, p. 14] defines an alarm as "an audible and/or visible means of indicating to the operator an equipment malfunction, process deviation, or abnormal condition requiring a response." Thus, the key point to distinguish such an alarm from nuisance alarms is on the requirement of operator response. Rothenberg [89] gives a brief definition of an alarm as "Alarm activation = Operator action." Therefore, in contrast to the nuisance alarms, a correct alarm is defined as the one that requires operators to pay attention or to take action in a prompt manner; otherwise, abnormal situations associated with correct alarms would have negative effects on operation safety and/or efficiency. Nuisance alarms are the major culprits for the phenomenon of alarm overloading. On the other hand, there are also the scenarios of having too many correct alarms, referred to as alarm floods (to be clarified later in Section III-D).

The consequences of alarm overloading are extremely detrimental to the important role played by alarm systems. First, a large number of alarms belong to the nuisance alarm group; they provide no useful information and only serve as distractions to plant operators. Due to "cry wolf" effect, nuisance alarms lead to confidence crisis of alarm systems. As a result, a correct alarm may be buried among such nuisance alarms and may consequently be overlooked by operators. Second, even if all occurred alarms are correct ones, e.g., those in alarm floods, the alarm rate may be too high to be manageable by operators. When the alarm rate is too high, operators have no choice but to ignore many of the occurred alarms. In this case, the designed functionality of alarm systems is completely discredited. As an example, two operators received 275 different alarms during the 10.7 min before the explosion accident occurred at the Texaco Refinery in Milford Haven [53].

III. Causes for Alarm Overloading

In order to alleviate the phenomenon of alarm overloading, the very first step is to find the main causes leading to such a phenomenon. This section identifies four main causes and provides industrial examples as supportive evidences.

A. Cause #1. Chattering Alarms Frequently Occur due to Noise/Disturbance

Chattering alarms are the mostly encountered nuisance alarms and may account for 10%–60% of alarm occurrences [89] (page 123). An analysis based on 75 alarm systems showed that on average over 70% of the alarm occurrences came from chattering alarms [52, p. 83]. The industrial standard ANSI/ISA-18.2 defines a chattering alarm as one that repeatedly transitions between the alarm state and the normal state in a short period of time [58, p.16]. As a result, there is no time or necessity for operators to analyze such alarms and take actions. Two closely related nuisance alarms are the fleeting and repeating alarms. Fleeting alarms also have short-time alarm duration, but do not immediately repeat [58, p. 74]. Repeating alarms are alarms rising and clearing repeatedly over a period of time [41, p. 95]. Chattering alarms are also named as cycling alarms [89, p. 444]. These types of alarms are typically generated due to random noise and/or disturbances on process variables configured with alarms, especially when the process variables are operating close to their alarm trippoints [41, p. 95]. In addition, chattering alarms can be induced by repeated on–off actions of control loops or regular oscillatory disturbances in process variables [105]; in this case, chattering alarms repeatedly make transitions between alarm and non-alarm states with regular (possibly large) time periods. To have a unifying terminology, all of these alarms are referred to as chattering alarms in this context, with a refined definition as: a chattering alarm is one that transitions between the alarm state and the normal state very quickly or with a constant time period.

Bransby and Jenkinson [17, App. 10] discussed some industrial examples of chattering alarms caused by the noise on a process variable that was operating close to the alarm trippoint, and by the oscillations from repeated on–off control actions having a regular oscillation period of 43 min. Hollifield and Habibi [52, p. 84] listed top 10 chattering alarms based on 150 days of data. Ahnlund *et al.* [5] showed chattering alarms associated with periodic signals or outliers in signals. Wang and Chen [105], [106] presented industrial examples of chattering alarms due to noise and oscillations at petro-chemical and thermal power plants. One example of chattering alarms from Plant A is presented next for illustration.

Example 1: A process variable $x(t)$ is the difference between the maximum and minimum values of measurements from 54 temperature sensors installed at stator outlet pipes at Plant A. $x(t)$ is configured with a high alarm variable $x_a(t)$ with an alarm trippoint at 8.0. That is, $x_a(t)$ takes the value 1 if $x(t)$ is

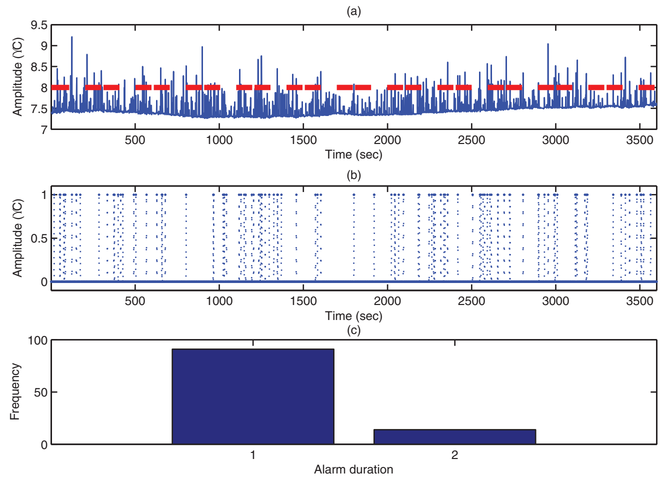


Fig. 5. An example of chattering alarms due to measurement noise at Plant A: (a) process variable (solid) and alarm trippoint (dash); (b) alarm variable; and (c) histogram of alarm durations.

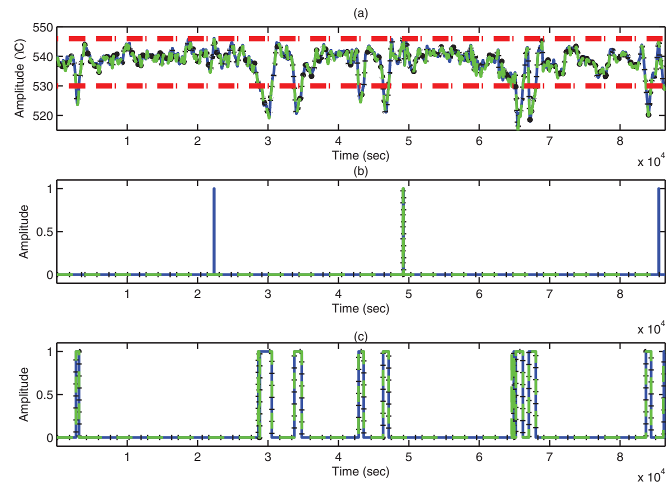


Fig. 6. (a) The measurements of main steam temperature from three sensors. (b) The samples of three high alarms variables. (c) The samples of three low alarms variables.

greater than 8.0, and the value 0, otherwise. Due to the measurement noise aggregation in 54 sensors, high-frequency noises contaminate $x(t)$ and lead to a large number of alarm occurrences, as shown in Fig. 5(a) and (b). Among the samples of $x(t)$ collected in 1 hour with sampling period 1 s, there were 105 alarm occurrences in Fig. 5(b). The alarm durations of these 105 alarms were all no larger than 2 samples, as shown in the histogram of alarm durations in Fig. 5(c). Here, the alarm duration is calculated as the number of consecutive samples taking the value 1 between each alarm occurrence and the corresponding clearance. Therefore, these occurred alarms are clearly nuisance alarms.

B. Cause #2. Alarm Variables Are Incorrectly Configured

Before the appearance of modern computerized monitoring systems such as the DCS and SCADA, each alarm variable was realized by hardware devices with high investment costs. As a result, each variable to be configured with alarm was carefully selected and thoroughly justified, and the total number of alarm variables was very limited, e.g., about 30 to 50 per process unit. As a comparison, alarm variables in modern computerized monitoring systems are very easily realized in a technical sense by clicking a mouse and entering alarm trippoint values at a computer, so that alarm configuration is regarded as "free" without any cost. In addition, configuring more alarm variables is often believed to be beneficial in improving operation safety. Hence, the number of alarm variables increases dramatically. For instance, Nimmo [85] reported that the number of alarm variables in one plant was increased from 150 for the hardware-based alarm variables to 14000 for the computerized alarm variables. Hollifield and Habibi [52, p. 13] indicated that the configured alarms per operator had increased exponentially from less than 100 alarm variables in 1960 to about 4000 alarm variables in 2000. Many variables are configured with alarms without a careful study on the necessity of configuring alarms and on the alarm priorities. As a result, there are a large number of variables that should not be configured with alarms or are configured with alarms in an incorrect manner. For instance, Timms [102] stated that some industrial facilities were with more than 15000 configured alarm variables, and over 50% of the configured alarm variables had been removed after alarm rationalization.

Example 2: The main steam temperature at Plant A is one of critical process variables, and is monitored by three independent sensors. The measurement of each sensor is configured with high-alarm trippoint 546 and low-alarm trippoint 530. Thus, there are three high (low) alarm variables associated with the main steam temperature. Fig. 6 presents the measurements of the three sensors and the corresponding six alarm variables in 24 hours. The high (low) alarm variables were very close for most of times; sometimes only one alarm variable ran into the alarm status for a short period of time due to noise effects, which certainly should not lead to operator's action such as adjusting the attemperator water flow in order to affect the main steam temperature. Clearly, it is not necessary to configure an individual high (low) alarm variable for each sensor; a more reasonable way is to design two alarm variables (one for the high alarm and another for the low alarm) by integrating the measurements from three sensors.

Example 3: There are more than 100 control loops at Plant A, playing critical roles in the plant operation. Each control loop is associated with two digital variables, namely, the manual mode variable and the switch-to-manual variable. Fig. 7 shows the historical data samples of the two variables for the feed air control loop at Plant A. The manual mode variable indicates whether the control loop is in the manual mode. For each control loop, the manual mode variable is configured with an alarm. Thus, when a control loop is switched from auto mode to manual mode, an alarm occurs. However, the switching from auto to manual mode is often done by operators. When an alarm occurs under such a circumstance, operators certainly take no actions to address the alarm. Therefore, there is no need to configure such a mode change with an alarm variable. By contrast, the switch-to-manual variable is the one forcing a control loop switching from the auto mode to manual, when some conditions are satisfied. It should be configured with an alarm under

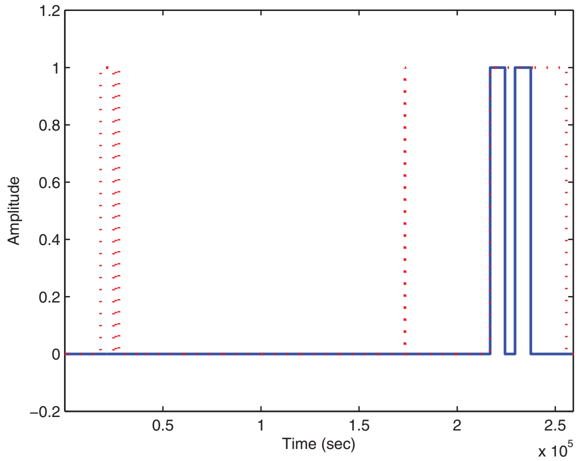


Fig. 7. The data samples of manual mode variable (dash) and the switch-to-manual variable (solid) for feed air control loop.

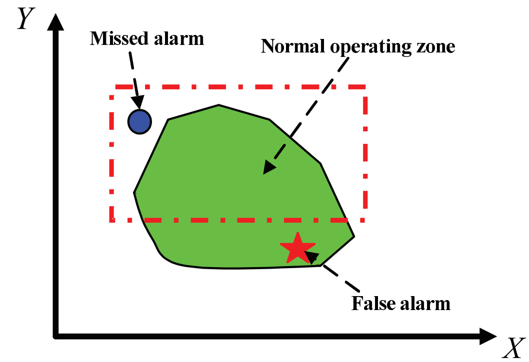


Fig. 8. Schematic diagram of normal operating zone with isolated alarm trippoints.

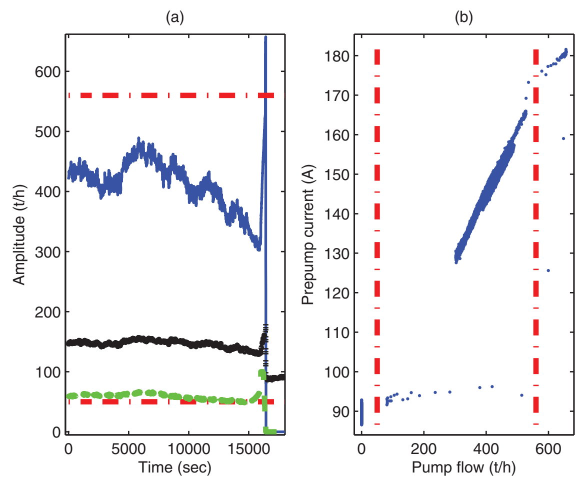


Fig. 9. (a) The time trends of inlet flow (blue solid), prepump current (black dotted), operator demand (green dashed) and high/low-alarm trippoints (red dash-dotted). (b) The scatter plot (dotted) of inlet flow and prepump current with high/low-alarm trippoints (red dash-dotted).

the condition that the control loop is in the auto mode. That is, when an alarm for the switch-to-manual variable occurs, operators need to check the corresponding conditions in order to return the control loop back to the auto mode.

C. Cause #3. Alarm Design Is Isolated From Related Variables

The main task in alarm design is to determine the mechanisms to generate alarm variables, and choose proper parameters of the mechanisms. For instance, the most common alarm variables are generated by comparing the continuous-valued measurements of analog variables with the high or low alarm trippoints; thus, the alarm trippoints are the design parameters to be chosen. However, in contemporary alarm systems, alarm design is usually isolated from other related variables, e.g., alarm trippoints are constant values and do not vary with other variables. It is a well-known fact that many process variables are related to each other via mass and energy conservation laws. Thus, a proper design of alarm generation mechanisms should be dependent on the related variables; otherwise, some types of nuisance alarms, including false alarms and missed alarms, may occur. As an illustration, Fig. 8 presents a schematic diagram of a normal operating zone of two correlated variables configured with alarms. If their alarm trippoints are designed in an isolated manner, then a rectangular area is formulated, being inconsistent with the normal operating zone. As a result, false alarms (missed alarms) are possibly present, shown as the star (circle) points in Fig. 8.

Example 4: The inlet flow of a feedwater pump at Plant A is configured with high-alarm trippoint 560 and low-alarm trippoint 50. In one abnormal situation, the drum water level was decreasing abruptly so that the operating demand for the feedwater pump was increasing quickly, shown as the dashed line in Fig. 9(a). As a result, the inlet flow [the solid line in Fig. 9(a))] exceeded the high-alarm trippoint 560, raising a high alarm. However, the increment of inlet flow was induced by operator's demand. That is, the feedwater pump performed normally as requested, instead of being at an abnormal condition as implied by the occurred high alarm. The prepump current [the dash-dotted line in Fig. 9(a)] was closely related to the inlet flow. As a matter of fact, the same proportional relation between the prepump current and the inlet flow was unchanged, even if the inlet flow exceeded the high-alarm trippoint 560, as shown in Fig. 9(b). Therefore, the occurred high alarm of the inlet flow was a false alarm due to the usage of constant alarm trippoint, which was isolated to the related variables such as the prepump current. A remedy is to design dynamic alarm trippoints for the inlet flow when the feedwater pump is running. The dynamic alarm trippoints are expected to be proportional to the prepump current, and form a normal operating band that could tolerate a certain level of uncertainties for such a proportional relation due to process variations.

After the high-alarm occurrence, another low alarm occurred, as shown in Fig. 9(a). This was owing to the shutdown of the feedwater pump, since the operating demand in Fig. 9(a) took the value of zero at the end. Such a low alarm did not say that the feedwater pump was in an abnormal condition, and no operator actions were needed. Therefore, the occurred low alarm was also a false alarm, which could be removed by incorporating the states of related variables such as the on/off state of the feedwater pump into the design of alarm variables, as described by the state-based alarming in ANSI/ISA-18.2 [58] and Hollifield and Habibi [52].

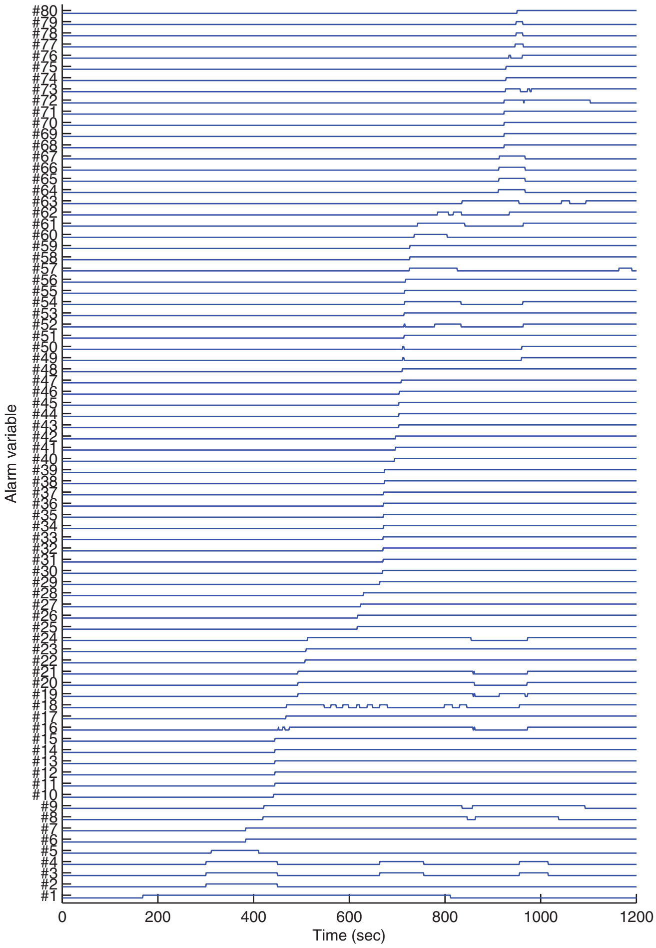


Fig. 10. Time trends of alarm variables in the alarm flood at Plant A on January 24, 2014 (the lower level for each alarm variable represents the value 0, and the higher level represents the value 1).

D. Cause #4. Abnormality Propagates Owing to Physical Connections

A large-scale industrial process is usually composed of upstream and downstream devices, which are physically connected. One abnormal condition in one process unit is very likely to be propagated to the downstream devices or the upstream devices owing to automatic control loops or recycling connections. As a result, a large number of correct alarms may arise in a short period of time for the process variables associated with these devices configured with alarms; these alarms appear to "flood" operators. Hence, this phenomenon is referred to as an alarm flood. In other words, an alarm flood is defined as the situation when the number of alarm activations exceeds the operator's ability to process them [89, p. 440]. There is no unique quantitative definition for alarm floods. A widely accepted definition is that an alarm flood begins when ten or more alarms occur within a 10-min time period until the alarm rate drops below five alarm occurrences in 10 min [89], [52], [58]. Rothenberg [89, p. 120] also defines serious alarm floods as the situations having no less than 100 alarms within 10 min, or having 10 consecutive time periods, where each time period has no less than 10 alarms within 10 min. Generally, the alarm floods include a large amount of nuisance

TABLE III
ALARM VARIABLES DURING THE ALARM FLOOD AT PLANT A ON JANUARY 24, 2014

#	Description	#	Description
1	Main steam pressure high	41	Mill B1 coal speed low
2	Mill C1 coal separator temperature #1 high	42	Mill B2 coal speed low
3	Coal system abnormal #1	43	Mill B2 #2 coal valve close
4	Coal system abnormal #2	44	Mill B1 #4 coal valve close
5	Mill C1 coal separator temperature #2 high	45	Mill B2 #1 coal valve close
6	MW control in manual	46	Mill B1 #2 coal valve close
7	MW control switch to manual	47	Mill B2 #3 coal valve close
8	Feed water A control error high	48	Mill B1 #3 coal valve close
9	Feedwater B control error high	49	Mill B2 seal/primary air pressure low
10	Feed air A in manual	50	Mill B1 seal/primary air pressure low
11	Feed air B in manual	51	Mill B1 coal separator temperature #2 high
12	O2 control in manual	52	Fire B1-2 off
13	Primary frequency regulation on	53	Fire B1-3 off
14	Boiler main control in manual	54	Fire B2-2 off
15	CCS mode off	55	Fire B2-3 off
16	Drum level #2 low	56	Mill B1 coal separator temperature #1 high
17	Turbine main control in manual	57	Mill B2 coal separator temperature #1 high
18	ATC invalid	58	Reheater temperature high
19	Drum level abnormal	59	Reheated steam temperature high
20	Drum level #3 low	60	Mill B2 coal separator temperature #2 high
21	Drum level low	61	Fire B1-4 off
22	Feedwater A control in manual	62	Fire B2-1 off
23	Feedwater B control in manual	63	Condensor flow high
24	Drum level #1 low	64	Drum level #1 high
25	Superheater temperature #1 high	65	Drum level #3 high
26	Superheater temperature #2 high	66	Drum level high
27	Superheater temperature #3 high	67	Drum level #2 high
28	Main steam temperature high	68	Feedwater A ETS trip
29	Mill C MTR	69	Feedwater A remote control off
30	Mill C1 #4 coal valve close	70	Feedwater A valve close
31	Mill C1 #3 coal valve close	71	Feedwater A emergency valve close
32	Mill C1 #1 coal valve close	72	Feedwater A/B speed error high
33	Mill C1 #2 coal valve close	73	Feedwater B inlet flow high
34	Fire C1-4 off	74	Feedwater A MEH trip
35	Fire C1-2 off	75	Feedwater A trip
36	Fire C1-3 off	76	Feedwater A inlet flow high
37	Fire C1-1 off	77	Drum level #2 high high
38	Mill C1 seal/primary air pressure low	78	Drum level #1 high high
39	Mill C2 seal/primary air pressure low	79	Drum level #3 high high
40	Mill B MTR	80	MFT

alarms. However, the true reason that "loss incidents frequently involved the operator being overloaded with alarm floods" [41, p. 139] is that too many correct alarms arise in a short time period. Taking the explosion accident at the Texaco refinery as an example, there were 275 alarms in the 10.7 min before the explosion [53]. Large numbers of correct alarms are usually due to the abnormality propagation. That is, a primary abnormal event results in consequential abnormal events; these events raise the related alarms [102]. Let us look at a specific example for a concrete view of alarm floods.

Example 5: One alarm flood occurred at Plant A on January 24, 2014. The alarm flood started with a high alarm of the main steam pressure (alarm tag #1 in Table III) that was raised at the time instant 23:36:07. However, the plant operators mistakenly adjusted the feed coal flow in a wrong direction, so that the main steam pressure kept increasing. The boiler water level decreased to an even lower position -298, which almost reached a low-low alarm trippoint -300. Then, the operator reduced the feed coal flow after realizing such a mistake by shutting down two coal grinding mills. The main steam pressure decreased dramatically; as a result, the boiler drum level increased too quickly to arrive at a high-high alarm trippoint at 300, which automatically triggered an emergency shutdown (alarm tag #80 in Table III) of the entire power generation unit at the time instant 23:49:13. During the 13 min from 23:36:07 to 23:49:13, 80 alarm variables ran into the alarm status, as shown in Fig. 10. The descriptions of the 80 alarm variables are given in Table III.

A retrospective investigation revealed that this alarm flood involved the propagation of several abnormalities. The time trends of several major process variables are presented in

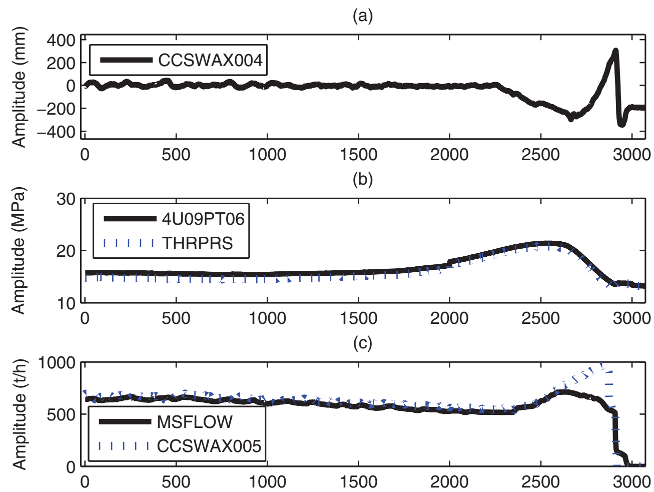


Fig. 11. Time trends of main process variables during the alarm flood at Plant A on January 24, 2014: (a) boiler drum level (tagname CCSWAX004) as the process variable associated with the alarm variables with Tags #16,19–21,24,64–67 and 77–79 in Table III; (b) steam pressure (tagname THRPRS) as the process variable associated with the alarm variable Tag #1, boiler drum pressure (tagname 4U09PT06); and (c) inlet water flow (tagname CCSWAX005) as the process variable associated with the alarm variables with Tags # 73 and 76, main steam flow (tagname MSFLOW).

Fig. 11. The abnormalities of the steam pressure (tagname THRPRS in Fig. 11) and the boiler drum level (tagname CCSWAX004 in Fig. 11) were the two most critical ones. Because the physical connection of the steam pipeline and the boiler drum, the boiler drum pressure (tagname 4U09PT06 in Fig. 11) was directly affected so that another abnormality appeared in the boiler drum due to the unbalanced inlet water flow (tagname CCSWAX005 in Fig. 11) and the main steam flow (tagname MSFLOW in Fig. 11); such an abnormality caused a large variation of the boiler drum level, leading to the emergency shutdown of the power plant. As some less critical abnormalities, the control loops of two steam-driven feedwater pumps were switched automatically into the manual mode due to the abnormal boiler drum level; in order to reduce the main steam pressure via manual operations, operators switched several control loops into the manual mode such as the feed air flow control loop, and adjusted the feed coal flow by switching off two coal grinding mills. The 80 alarms associated with all abnormalities were raised in 13 min, resulting in an alarm flood. Apparently, the operators failed in diagnosing and handling the abnormalities, so that the abnormalities led to the incident of process unit emergency shut down. It is worthy to point out that the alarm occurred first in time does not always indicate the origin of abnormalities, because the order of occurrences depends on several factors such as the configuration of alarm trippoints and the variation speeds of process variables.

Remark #1: The above-mentioned four causes could be mixed or present simultaneously in practice. For instance, the alarm floods in Examples 5 also involved the main cause #2, because Tags # 25–27 (superheater temperature high alarm variables) are similar to those in Example 2, and should be integrated to configure one alarm variable, instead of the three individual alarm variables.

IV. Current Research Status

This section summarizes the current research status for industrial alarm systems. The summary focuses on the state of the art of methodologies related to the four main causes given in Section III that lead to the phenomenon of alarm overloading. Hence, the cited references are not going to represent a complete list of publications on industrial alarm systems; see, e.g., Kirschen and Wollenberg [66] and Kim [65] for a review of earlier methodologies for alarm systems and diagnostic systems in the power industry.

A. State of the Art in Dealing With Chattering Alarms

The first step in dealing with chattering alarms is to detect the presence of chattering alarms. Yuki [125] and Noda *et al.* [86] detected chattering or unnecessary alarms by focusing on the balance between alarm occurrences and operator actions. Kondaveeti *et al.* [68] quantified the degree of chattering alarms based on the alarm run lengths. Naghoosi *et al.* [82] estimated the chattering index based on statistical properties of process variables. Wang and Chen [105] revised the chattering index by taking the number of data samples into consideration. Instead of using chattering indices, Wang and Chen [106] formulated two hard rules to detect chattering alarms.

The second step in handling chattering alarms is to design alarm systems to reduce the number of chattering alarms in the future. Burnell and Dicken [22] introduced an auto-shelving mechnasium to deal with repeating alarms. Bransby and Jenkinson [17, App. 10] and the guideline EEMUA-191 [41, App. 9] recommended filters, deadbands, delay timers, and shelving mechanisms to handle repeating alarms. Ahnlund *et al.* [5] separated process variables into 14 classes and selected proper filters accordindg to the classes to deal with chattering alarms. Henningsen and Kemmerer [50] and Srinivasan *et al.* [93] temporarily changed alarm trippoints or put alarm variables in a shelving condition based on statistical process control techniques. Hwang *et al.* [56] applied control charts to design a pre-alarm system to reduce the alarm frequency. Hugo [55] designed adaptive alarm deadbands to reduce the number of chattering alarms. Naghoosi *et al.* [82] designed optimal alarm limits and deadbands by minimizing the chattering index. Wang and Chen [105], [106] proposed online methods to remove chattering and repeating alarms by adjusting alarm trippoints or using delay timers.

B. State of the Art in Handling Incorrectly Configured Alarm Variables

In terms of configuring process variables with alarms, there are mainly two types of alarm configuration methods. The first type is to establish relations between process variables and abnormal situations based on *process knowledge* in order to configure alarm variables accordingly. Yan *et al.* [115] drew up an abnormality propagation diagram and selected suitable alarm variables for malfunctions. Yan *et al.*[116] allocated sensors for alarm variables to optimize fault detection reliability based on process topology represented by signed directed graphs. Takeda *et al.* [95], [96] selected process variables to be configured with alarms based on cause-effect models from process topology. Dalapatu *et al.* [34] assigned alarm variables to a group of process variables based on process knowledge to identify abnormal events.

Another type of alarm configuration methods is to analyze correlated or consequential alarm variables based on *historical data*, to remove redundant alarm variables or redesigning alarm variables. Geng *et al.* [46] and Zhu and Geng [127] grouped alarm variables into different clusters and ranked the priorities of alarm variables in each cluster via a fuzzy clustering-ranking algorithm. Kordic *et al.* [70] found correlated sets of alarm variables by using time intervals between alarm occurrence and disappearance time instants. Kondaveeti *et al.* [69] built an alarm similarity map based on the Jaccard similarity index for binary-valued alarm variables. Yang *et al.* [118] clustered correlated alarm variables based on correlation coefficients of pseudo continuous time series generated from binary alarm data samples. Yang *et al.* [120] exploited the Sorgenfrei similarity coefficient for binary-valued alarm variables and the distribution of correlation delays to detect correlated alarm variables. Wang *et al.* [108] found consequential alarms based on correlation delays and Granger causality measures.

In terms of assigning different priorities to alarm variables, qualitative approaches are the common practice. For instance, Timms [101] discussed alarm prioritization based on classifying the consequences on personnel safety, financial loss and environmental consequences into several categories. Very few quantitative approaches appeared recently. Chang *et al.* [25] proposed a quantitative risk-based approach to prioritize alarm variables by integrating the process safety time together with the probability and impact of potential hazards. Arifin and Choudhury [8] quantified the system failure probability to sort the importance of alarm variables.

C. State of the Art in Alarm Design by Considering Related Variables

As stated in Section III-C, the design of the generation mechanisms for alarm variables often needs to take correlated variables into consideration. Since multiple process and alarm variables are involved, the alarm generation mechanisms are diverse, developed without a unifying framework. Yamanaka and Nishiya [114], Charbonnier *et al.* [26], and Charbonnier and Poret [27] extracted the trends of several related process variables to formulate different episodes for generating alarm variables. Bristol [18] suggested automatic adaptation of alarm trippoints to varying process situations. Brooks *et al.* [19] proposed a geometric process control method to obtain dynamic alarm trippoints from multivariate best operating zones. Nihlwing and Kaarstad [84] developed a state-based alarm system that dynamically presented alarms based on 19 different process states for a nuclear power plant simulator. Jang *et al.* [61] developed different rules to filter or suppress alarms for a nuclear power plant reactor. Izadi *et al.* [60] and Kondaveeti *et al.* [67] applied multivariate statistics to generate alarms more efficiently. Yang *et al.* [117] performed correlation analysis for alarm signals to optimize alarm trippoints. ANSI/ISA-18.2 [58], Hollifield and Habibi [52], Arjomandi and Salahshoor [9], and Jerhotova *et al.* [62] described state-based alarming to avoid long-standing alarms by changing alarm settings according to process states. Gupta *et al.* [49] exploited wavelet analysis and principal component analysis to alleviate noise effects and detect faults in pharmaceutical manufacturing processes. Zhu *et al.* [128] obtained dynamic alarm trippoints depending upon multiple steady states and transitions between these states. Zang and Li [126] optimized alarm trippoints by minimizing FAR and MAR based on joint probability density of multiple process variables. Alrowaie *et al.* [6] proposed a model-based alarm design method based on particle filtering for multivariate nonlinear stochastic systems. Xu *et al.* [113] predicted the impending alarms based on hybrid models utilizing both first principles and data, so that operators could have more time to handle alarms.

D. State of the Art in Processing Alarms Owing to Abnormality Propagation

Abnormality propagation is a major reason for multiple alarms and even alarm floods. Some techniques have been developed to deal with this case in alarm floods. Laberge *et al.* [71] addressed alarm floods through a new alarm summary display design by showing alarms in a time series represented by icons together with short alarm descriptions. Varga *et al.* [104] detected the development of unsafe situations and suggested the operators with necessary safety actions based on the prediction of physical models. Basu *et al.* [12] ordered alarms on the power grid according to different severity measures. Tchamova and Dezert [97] estimated the degree of danger for alarms based on Dezert-Smarandache theory to fuse conflicting evidences. Blaauwgeers *et al.* [16] and Zhu *et al.* [129] used the BowTie diagram and Bayesian network respectively, to perform real-time risk analysis for operators to prioritize alarm handling.

To reduce the number of alarms due to abnormality propagation, suppression of consequential alarms or redesign of alarm variables are recommended, in particular, during alarm floods. The guideline in EEMUA 191 [41] suggested reviewing consequential alarms and using alarm grouping to reduce the number of alarm activations during alarm floods. Hollender and Beuthel [51] suggested hiding consequential alarms based on causal relations between process variables. Beebe, Ferrer and Logerot [15] suggested state-based alarm rationalization to control alarm floods whose occurrences are typically due to a change of process states.

Obviously, the essential solution to consequential alarms owing to abnormality propagation is to find one or more abnormalities as the root cause(s) of a set of occurred alarms. The first type of methods to find the root causes is based on the process knowledge and/or learning algorithms. Young *et al.* [123], McDonald *et al.* [80] *et al.* [109], and Wen *et al.* [110], respectively, used a Tabu-search method and a refined genetic algorithm to find the abnormal events to explain a given set of reported alarms from knowledge-based tables describing the relations between abnormal events and alarm variables for power systems. Cauvin *et al.* [23] used causal graphs and models to interpret the root causes of alarms. Dashlstrand [35], Souza *et al.* [92], Larsson *et al.* [73], and Tolga *et al.* [103] introduced multilevel flow models or fuzzy neural networks to analyze root causes of alarms. Kezunovic and Guan [64] used fuzzy reasoning Petri-nets techniques to diagnose the root cause of alarms. Liu *et al.* [76] introduced an operator model as a virtual subject to evaluate the performance in diagnosing root causes of alarms. Simeu-Abazi *et al.* [91] exploited dynamic fault trees to locate faults from alarms with application to avionic systems. Guo *et al.* [48] and Wei *et al.* [111] determined fault/disturbance causes based on rule networks or temporal constraint networks between cause hypothesis and alarms for digital power substations. Abele *et al.* [1] and Wang *et al.* [107] exploited Bayesian networks to analyze the root cause of alarms in an online manner. Dubois *et al.* [37] and Lee *et al.* [74] adopted logic diagrams to perform real-time cause-effect analysis for alarms. The above methods need rather complete and accurate knowledge about the process, which is sometimes difficult to obtain in practice, especially when the process is large and complex.

Alternatively, historical alarm series can be used to extract time patterns of alarms. For this purpose, sequential pattern matching has been introduced for examining alarm series. Folmer *et al.* [43], Folmer and Vogel-Heuser [44] and Folmer *et al.* [45] used this idea to cluster frequent occurring subsequences in alarm logs and identify alarms with causal relations to redesign alarm systems for reducing the number of alarm variables. Similar historical alarm floods can be exploited to extract representative information. Ahmed *et al.* [4] located similar alarm floods based on the consecutive alarm frequencies and used dynamic time warping to obtain optimal matching between two alarm floods. Cheng *et al.* [30] proposed a modified Smith–Waterman algorithm for local alignment of two alarm flood sequences so that common alarm sequence segments could be extracted. Charbonnier *et al.* [28] extracted the fault sequence template from alignments of alarm sequences from the same fault, and compared a new alarm sequence to the template for fault isolation. This type of methods rely on sufficient historical alarm data, including all possible patterns. However, we may not encounter all the abnormalities in the past, leading to its failure for new patterns in online practice.

Another type of methods is to capture the plant topology in advance to describe the intrinsic structure of the process. When the process enters an abnormality, a backtrack or hypothesis test can be employed based on the current symptom to find the root causes. There are plenty of process data and various methods that can be used for capturing this causal topology. Bauer *et al.* [14] used time-delayed cross correlation to identify the propagation paths and then build a causal map. Similarly, a series of causality identification methods have been proposed, such as Granger causality [124], partial directed coherence [47], transfer entropy [13], direct transfer entropy [38], and transfer zero-entropy [40]. Nonlinearity can also be used as an indicator in causality analysis, e.g., Thornhill [100] used a nonlinearity index to find the root cause because nonlinearity is strongest at its source. Cecilio *et al.* [24] used the nearest neighbors method to identify abnormality in each series and showed the abnormality propagation order as a color plot. By the total contribution plot, one can observe how a fault is spreading across the process [11]. For a survey refer to [121] and [39]. The limitations of these data-based methods are assumptions of model linearity or data stationarity, and the high computational burden. In addition, these methods are usually used for obtaining a short list of possible root causes that cannot be confirmed without resorting to process knowledge. Therefore, in addition to process data, another important resource for building plant topology is plant connectivity information that describes physical, mass or information linkage between process units. Adjacency matrices [63], signed directed graphs [78], XML description [122] and semantic description [121] are efficient representations of plant connectivity, which provide a physical foundation of plant topology with extra and redundant information. These connectivity-based methods should be integrated with data-based methods for validation. Thambirajah *et al.* [98] combined the cause-and-effect matrix derived from measurements and qualitative information about the process layout. Landman *et al.* [72] used a dedicated search algorithm to validate the quantitative results of the data-driven causality using the qualitative information on plant connectivity. Yang *et al.* [119] validated the data- and connectivity-based results mutually. Alarm data can also be used, e.g., Schleburg *et al.* [90] combined plant connectivity and alarm logs.

The combination of different resources and methods is a proper choice in real applications according to availability of resources and application objectives. Chiang and Braatz [32] integrated the statistical analysis with the causal map. Thornhill *et al.* [99] enhanced data-based analysis process understanding. Maurya [79] combined signed directed graphs with qualitative trend analysis. Di Geronimo Gil *et al.* [36] merged first-principles structural models with plant topology derived from a process drawing.

V. Research Problems to be Solved

In this section, we first formulate the lifecycle of alarm variables into three stages, namely, alarm configuration, alarm design, and alarm removal, and connect the stages with the four main causes in Section III. Next, for each stage, we propose some fundamental research problems to be solved, on the basis of research status summarized in Section IV.

A. Lifecycle of an Alarm Variable

The industrial standard ANSI/ISA-18.2 [58] and the guideline in EEMUA-191 [41] defined desired performance benchmarks for industrial alarm systems, e.g., the second column in Table II. In order to achieve the benchmarks, the industrial standard ANSI/ISA-18.2 presented ten stages for an alarm management lifecycle, namely, alarm philosophy, identification, rationalization, detailed design, implementation, operation, maintenance, monitoring and assessment, management of change, and audit [58, p. 22]. Hollifield and Habibi [52] listed seven steps to achieve a highly effective alarm system:

Step 1) Develop, adopt and maintain alarm philosophy.
Step 2) Benchmark alarm system.
Step 3) Find bad actor alarms.
Step 4) Perform alarm documentation and rationalization.
Step 5) Implement alarm audit and enforcement technology.
Step 6) Implement real-time alarm management.
Step 7) Control and maintain improved alarm system.

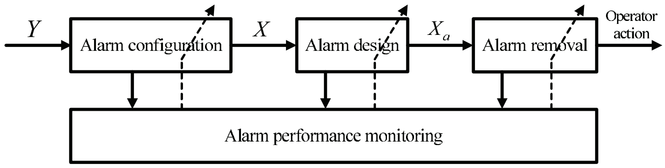


Fig. 12. The lifecycle of alarm variables.

These stages and steps have been proven to be effective pathways to follow in some practical applications, but more techniques need to be developed to support these stages and steps. The techniques, as the expected outcomes of the fundamental research problems to be proposed in the next subsections, are better to be presented in a different way with respect to the alarm variables.

The lifecycle of alarm variables may be formulated as three stages, namely, alarm configuration, alarm design, and alarm removal, as shown in Fig. 12. That is, some process variables in the set Y are selected in the stage of alarm configuration to formulate the set X as the process variables to be configured with alarms, X_a is the set of alarm variables based on the alarm design for X, and operator action needs to be decided in the stage of alarm removal once X_a is in the alarm state. In particular, each stage has its own alarm performance monitoring techniques in order to have a systematic solution to the problems therein and to maintain a satisfactory performance of each stage in a long run. The three stages are closely related to the four main causes discussed in Section III. The second cause "alarm variables are incorrectly configured" is about alarm configuration; the first and third causes are for nuisance alarms, which are the main concerns in the alarm design; the fourth cause "abnormality propagates owing to physical connections" is related to the alarm removal, since one of key steps in alarm removal is to find the root causes of alarms, especially during an alarm flood. The three stages in Fig. 12 are also connected with the ten stages in [58] and the seven steps in [52], e.g., the alarm configuration are in line with steps 1 and 4, the alarm design are related to steps 2, 3, 5 and 7, and the alarm removal encloses step 6.

B. Research Problems in Alarm Configuration

Alarm configuration mainly involves the selection of process variables to be configured with alarms, and the determination of alarm priorities. There are three fundamental research problems to be solved:

Problem 1: Should an alarm variable be configured for a given process variable?

Problem 2: Which priority should an alarm variable be assigned?

Problem 3: Are there any incorrectly configured alarm variables in an existing alarm system?

In terms of Problem 1, the main concern is the determination of relations between abnormal events and process variables so that proper process variables are selected to identify abnormal events. Mathematically, the relation can be described as

$$X = g(Y, A) \tag{2}$$

where A stands for the set of abnormal events, Y is the set of candidate process variables, and X is the set of selected process variables to be configured with alarms. If the relation between abnormal events and process variables is established, then the process variables, whose variations reveal the presence of abnormalities, are selected to be configured with alarms. However, the existing approaches [115], [116], [95], [96], [34] summarized in Section IV-B to establish such a relation are based on process knowledge or obtained from historical operating data. Either way has its own limitations, e.g., the process knowledge is often incomplete and qualitative, while the data-based approaches are confined by the occurred phenomena in the historical data sets and contaminated by noise/disturbances. Hence, the main challenge is to exploit process knowledge and historical data *simultaneously* in establishing relations and selecting proper process variables to be configured with alarms.

In terms of Problem 2, the current approaches in industrial practice are mostly qualitative, based on the understanding of severity in missing alarms and a rough estimation of safety time in handling alarms. Quantitative approaches are the ones to be developed as alternatives or complements to the qualitative approaches. For instance, a more accurate estimation of the safety time could be obtained from historical data sets. Here, the safety time is the time period allowed to handle an alarm so that the associated negative consequences do not occur. The risk-based approaches in [8] and [25] are promising quantitative approaches to assign priorities for alarm variables. Alarm priorities may not be static; instead, dynamic alarm priorities are perhaps more reasonable for different process states and operational scenarios, as shown by [12], [16], [97], [129].

In terms of Problem 3, the incorrectly configured alarm variables may present themselves in different forms. The redundant (duplicated) alarm variables are the ones that always run into the alarm status simultaneously or in a constant time delay. The redundant alarm variables yield no extra useful information and should not be configured with alarms. The correlated and consequential alarm variables may also be resulted from incorrect alarm configurations, e.g., the highly correlated alarm variables from three sensors for the main steam temperature in Example 2. The correlated alarm analysis methods [69], [108], [118], [120] listed in Section IV-B are able to effectively detect the redundant, correlated and consequential alarm variables. If operator actions cannot be clearly defined for the occurred alarms, then the corresponding alarm variables may be incorrectly configured, as shown by the alarms from the switch-to-manual variable in Example 3. Yuki [125] and Noda *et al.* [86] associated operator actions with occurred alarms; if no operator actions could be found, then the occurred alarms were nuisance. However, the challenges for doing so is that many operator actions or responses are not recorded in the historical database, and even for the recorded ones, they are difficult to be associated with occurred alarms in an automatic manner.

C. Research Problems in Alarm Design

The first objective in the stage of alarm design is to design an alarm generation mechanism that transforms process variables X to be configured with alarms into alarm variables X_a as shown in Fig. 12, i.e.,

$$X_a = h(X, X_r; \theta).$$

Here, X_r stands for the set of process variables related to X, and θ is the vector of design parameters in the alarm generation mechanism $h(\cdot;\theta)$. The alarm generation mechanism $h(\cdot;\theta)$ includes the simplest one in (1), alarm deadband (hysteresis), delay timers and filters, which have been commonly adopted in practice, as well as more complex ones such as logic/model/knowledge-based alarming methods [41], [52], [58], [89]. Thus, the corresponding fundamental research problem to be solved is:

Problem 4: *How to design an alarm generation mechanism with good performance?*

One main objective in the design of $h(\cdot;\theta)$ is to reduce the number of nuisance alarms caused by noise and/or disturbance (main cause #1 in Section III); another main objective is to take the related variables into the generation of alarm variables (main cause #3). In particular, if related variables have to be considered, the alarm generation mechanisms have many different forms, as shown by the references cited in Section IV-C. However, there are two main challenges, which have not been well addressed in the existing studies, in order to solve Problem 4 in a systematic manner.

First, the normal and/or abnormal operating zones for process variables to be configured with alarms are difficult to obtain. The normal and/or abnormal operating zones are the prerequisites for the design of $h(\cdot;\theta)$. Some physical principles and laws may be exploited to build up mathematical models for process variables in the normal or abnormal conditions. However, the uncertainties of model parameters have to be estimated from historical normal/abnormal data sets to formulate the operating zones, instead of trajectories, to tolerate the variations of normal operations and the effects from noise/disturbances. The operating zones could also be established directly from the normal and abnormal historical data sets. However, these data sets are often not available at hand, and have to be separated from historical data samples. One argument is to do the classification manually via visualization and consultation to plant operators, which is possible only for small sized data sets. Few existing studies have addressed this research challenge. Xu *et al.* [112] proposed a new method to detect the data sets with different sample means and to compare the sample means to alarm trippoints via hypothesis tests, in order to isolate normal and abnormal data sets of one single process variable. If multiple process variables are involved, the static alarm trippoints cannot truly reflect the normal operating zone, as shown in Fig. 8. As implied by the industrial example shown in Fig. 9 and several [26], [27], [114] cited in Section IV-C, abnormalities may be detected by monitoring the consistency of changing directions of X and X_r. By incorporating this process knowledge, change detection methods and data clustering techniques need to be developed in the context of industrial alarm systems.

Second, the relation between $h(\cdot;\theta)$ and a performance index vector η is difficult to establish. The false alarm rate (FAR) and missed alarm rate (MAR) are the most commonly-adopted performance indices. It is a well-known fact that there is a tradeoff between the FAR and MAR. Thus, a loss function $l(\eta)$ can be formulated to balance the conflicting indices, e.g.,

$$l(\eta) = w\mathrm{FAR} + (1-w)\mathrm{MAR}$$

where the real number $w \in [0,1]$ is a weighting factor. An optimal design of $h(\cdot;\theta)$ is obtained by minimizing $l(\eta)$, i.e.,

$$h_{\mathrm{opt}}(\cdot,\theta_{\mathrm{opt}}) = \arg\min_{h(\cdot,\theta)} l(\eta).$$

One critical knowledge required in the optimization is the quantitative relation between $l(\eta)$ and $h(\cdot;\theta)$. For certain univariate alarm variables where the one-dimensional process variable X is independent and identically distributed (IID), this relation has been established for alarm deadband, delay timers and filters [2], [3], [29], [112]. However, the relation is rather difficult to establish under more practical assumptions, e.g., $x(t)$ is not IID. As an attempt, Alrowaie *et al.* [6] adopted particle filtering to estimate FAR and MAR for non-IID process variables from nonlinear stochastic systems. If X are multivariate and X_r are involved, the relation is much harder to be obtained. As a result, many existing studies in Section IV-C were limited to the proposition of $h(\cdot;\theta)$ without theoretical analysis on the performance, but only with some examples showing the empirical effectiveness.

The second objective in the stage of alarm design is to detect the presence of nuisance alarms for industrial alarm systems being in service, in order to initiate a redesign of $h(\cdot;\theta)$ as described in Problem 4 to reduce the number of nuisance alarms in the future. Thus, the corresponding fundamental research problem is:

Problem 5: Are there too many nuisance alarms, so that alarm generation mechanisms need to be redesigned?

Clearly, a solution of isolating historical normal and abnormal data sets partially solves Problem 5, because the availability of normal data segments directly classifies the occurred alarms in the normal data segments as false alarms. However, such a solution may be difficult or costly to be obtained, it would be desirable to attack Problem 5 via different approaches.

One approach to solve Problem 5 is based on some special characteristics of nuisance alarms. As an example, the rationale of the methods in [105] and [106] to detect chattering and repeating alarms is to look at the statistical regularity of alarm durations or intervals in historical alarm data samples. The main challenge in doing so is from the diverse types of nuisance alarms. Each type of nuisance alarms has its own characteristics to be exploited for detection. Some characteristics are easier to be captured. For instance, chattering and stale alarms are the alarms whose time durations are very short (e.g., less than 20 s) [106] or exceptionally long (e.g., more than 24 hours) [41], [58]. By contrast, the characteristics of some nuisance alarms are hard to be described, e.g., those due to the third main cause in Section III.

Another approach is to associate operator actions with occurred alarms [86], [125]. According to the definition of correct alarms in Section II, if an operator action is associated with the occurrence of an alarm, then the occurred alarm is correct; otherwise, it belong to a nuisance alarm. However, such an approach may not be feasible in practice, as commented earlier for Problem 3.

A more feasible approach to solve Problem 5 is to connect the occurred alarms with their consequences. If no harmful consequences have been detected, then the occurred alarms are certainly nuisance. The relation between alarm variables and their consequences is the one to be established in the stage of alarm configuration. Owing to the presence of noises or disturbances, the detection of consequences may rely upon some hypothesis tests to make a statistical classification.

After the detection of nuisance alarms, the severity of nuisance alarms needs to be evaluated to determine whether a redesign of alarm systems is necessary or which alarm variables need to be addressed with high priorities. For instance, the chattering indices in [68], [82], and [105] are used for this purpose.

The third objective in the stage of alarm design is to generate a predictive alarm to indicate upcoming critical abnormalities, so that operators could have more time to analyze the upcoming alarms and take proactive actions. Thus, the research problem is:

Problem 6: How to design mechanisms to generate predictive alarms in order to forecast upcoming critical abnormal events?

A standard approach is to generate predictive alarms based on time series modeling and prediction techniques, e.g.,

$$\hat{X}(t+i) = p(X(t), X(t-1), \ldots)$$

where $\hat{X}(t+i)$ with $i > 1$ is the prediction of a process variable $X(t)$ configured with alarms, from a predictor $p(\cdot)$ based on the current and past data sample $X(t), X(t-1), \cdots$. Such an approach may not work well due to complexity of process variables in practice.

A special attention has to be paid to alarm floods, which should be avoided as much as possible as due to the equalities "floods $=$ incidents $=$ loss" [15]. Hence, it is important to predict the upcoming of alarm floods and take preventive actions to avoid the occurrence of alarm floods. One approach for the prediction of alarm floods is based on the physical or hybrid models, which predict the evolution of process variables [104], [113]. However, developing physical models are technically challenging and time consuming. Hence, such an approach is only feasible and worthwhile for a limited number of critical devices or equipments, not applicable to general alarm variables. Since alarm floods are usually composed by alarm variables having physical connections, as shown in Section III-D, historical alarm floods may have certain regularities to be exploited. For instance, the switching-off events of coal grinding mills in Example 5 in Section III-D always lead to the occurrences of the fire-off alarms and the mill sear/primary air pressure low alarms. Thus, another approach to handle alarm floods is to detect similar historical alarm floods, extract regular patterns of these similar alarm floods, and predict an upcoming alarm flood by comparing the currently occurred alarms to the regular patterns. The related methods in [4], [28], and [30] cited in Section IV-D detected similar alarm floods and their regular patterns solely based on historical data, whose validation is rather difficult. A more convincing conclusion may be obtained by complementing these methods with the physical connections of alarm variables. Hence, the main challenge is to obtain the process knowledge related to similar alarm floods, and transform it into a form that could be incorporated together with the statistical regularity from historical data. Another challenge is about the computational speed of algorithms in the predication of alarm floods. The modified Smith–Waterman algorithm in [30] is perhaps the state of the art for alignment of similar alarm floods; however, the algorithm is only suitable for offline usage, due to the slow computation speed, especially for alarm floods with long sequences. Hence, the computation has to be improved greatly, in order to be fast enough for matching an occurring alarm flood to its similar ones in the historical database in an online manner.

D. Research Problems in Alarm Removal

Alarm removal is mainly concerned with analyzing root causes leading to the occurrence of alarms, and advising operators to take some proper actions to avoid the deterioration of negative consequences associated with alarms, and drive the process variables back to their normal operating zones so that the occurred alarms are eventually removed. There are two fundamental research problems to be solved in this stage:

Problem 7: What are the root causes of the occurring alarms?

Problem 8: What actions should operators take to address the occurring alarms?

The very first step in solving Problem 7 is to tell whether the occurring alarms have meaningful root causes, or in other words, they belong to nuisance alarms or correct ones. Nuisance alarms require no operator action or response, and their removal is one of the main objectives for the stages of alarm design. If an alarm variable always produces nuisance alarms, then the removal of these nuisance alarms is done by removing the configuration of the alarm variable. A hard online classification of nuisance and correct alarms is rather difficult, but a classification using statistical inference is feasible. For instance, a hypothesis test can be formulated with the following null and alternative hypotheses

$$H_0: X_a(t) \in \text{nuisance alarms};$$
$$H_1: X_a(t) \notin \text{nuisance alarms}.$$

Such a hypothesis test can be based on the statistical characters of alarm and/or process variables revealed in historical data sets, e.g., the probability mass functions of time durations of alarm variables.

For correct alarms, the objective of Problem 7 is to find out the occurring abnormalities as the root causes of alarms. As pointed out in Example 5, the alarm occurred first in time does not always indicate the origin of abnormalities, which certainly complicates the root cause analysis. If the relation between the abnormalities and alarm variables can be established in some manner, then the root causes could be found backwards. That is, the root causes of occurring alarms are located based on the inverse model of the relation $g(\cdot)$ in (2), i.e.,

$$\hat{A} = g^{-1}(X).$$

Here, $\hat{A}$ is the estimate of root causes for the alarms in X. There are a large variety of representations of $g(\cdot)$ and the approaches to yield $\hat{A}$; see the references cited in Section IV-D. In this sense, Problems 1 and 7 share a common objective to establish the relation $g(\cdot)$ between process variables and abnormalities. Therefore, as commented for Problem 1, the limitations of process knowledge and historical operating data are the main challenges in solving Problem 7, too.

Due to the large scale and complexity of industrial processes, the above mapping in (2) is more related to plant topology than statistical models. It is more fundamental to locate the root causes for an event or an alarm flood. As discussed in Section IV-D, nonlinearity, as a local metric, can be used to find the most possible root causes. To obtain the abnormality propagation, causality capturing methods are necessary based on process measurements as well as alarm series. All the nonlinearity- and causality-based methods need sufficient historical data. However, for a new process without historical data, or a process that do not have so much trouble in the past, it is impractical to obtain sufficient statistical data in all abnormal situations. As a result, plant connectivity should be taken into account to describe the interior relationship between process units and process variables. In practice, the information in the historical data and the plant connectivity should be integrated to improve their efficiency and accuracy.

The objective of Problem 8 is to provide operators quantitative or qualitative advices in taking proper actions to eventually remove the occurring alarms. Brooks *et al.* [19] claimed that corrective changes of manipulated variables could be advised from the geometric process control method based on best operating zones in a multivariate framework, and these advices would be valuable to help operators to take proper actions; however, no technical details on the parallel coordinate techniques and the projective geometry theory were revealed in [19]. These advices are indispensable, especially when many process and alarm variables are involved. The multiple occurring alarms have to be ordered on the basis of their severities of consequences, in order not to deviate further from the normal operating zones and lead to aggravation or even incidents [12], [16], [104], [129]. This is closely related to Problem 2, where the main challenge is to develop quantitative approaches along with the ones solely based on the process knowledge. In terms of alarm floods, the best way perhaps is to take preventative actions in order to avoid the occurrence of alarm floods, on the basis of the alarm flood prediction that has been discussed for Problem 6.

VI. Conclusion

Industrial alarm systems are receiving increasing attention from both industrial and academic communities. This is essentially owing to a gap between two facts that alarm systems are critically important for plant safety and efficiency on one hand, but on the other hand, alarm systems are suffering from poor performance of having too many alarms to be handled by operators. A necessary step to resolve such a gap is to find out the main causes for the phenomenon of alarm overloading. This paper attempted to do so by identifying the four main causes, namely, chattering alarms due to noise and disturbance, alarm variables incorrectly configured, alarm design isolated from related variables, and abnormality propagation owing to physical connections. The literature survey in Section IV showed the different maturity levels of existing methodologies in addressing the four main causes. For instance, some recent studies cited in Section IV-A have made promising progress in the detection of chattering alarms and the systematic design of alarm generation mechanisms to reduce the number of chattering alarm occurrences, while the study on alarm floods is still in its infancy with very few published results. Eight fundamental research problems to be solved were presented for the lifecycle of alarm variables composed of three stages, namely, alarm configuration, alarm design and alarm removal. As these problems originate from industrial practices, it would be crucial to investigate them by analyzing industrial data samples of process and alarm variables, together with process knowledge, and to validate the solutions in real-time industrial applications.

Acknowledgment

The authors would also like to thank Shandong Electric Power Research Institute in China for providing industrial data of Plant A, and research partners for sharing their understanding on industrial alarm systems.

References

[1] L. Abele, M. Anic, T. Gutmann, J. Folmer, M. Kleinsteuber, and B. Vogel-Heuser, “Combining knowledge modeling and machine learning for alarm root cause analysis,” in *Proc. 7th IFAC Conf. Manuaf. Modeling, Manage. Control*, Saint Petersburg, Russia, Jun. 19–21, 2013, pp. 1843–1848.

[2] N. A. Adnan, I. Izadi, and T. Chen, “On expected detection delays for alarm systems with deadbands and delay-timers,” *J. Process Control*, vol. 21, pp. 1318–1331, 2011.

[3] N. A. Adnan, Y. Cheng, I. Izadi, and T. Chen, “Study of generalized delay-timers in alarm configuration,” *J. Process Control*, vol. 23, pp. 382–395, 2013.

[4] K. Ahmed, I. Izadi, T. Chen, D. Joe, and T. Burton, “Similarity analysis of industrial alarm flood data,” *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 2, pp. 452–457, Apr. 2013.

[5] J. Ahnlund, T. Bergquist, and L. Spaanenburg, “Rule-based reduction of alarm signals in industrial control,” *J. Intell. Fuzzy Syst.*, vol. 14, pp. 73–84, 2003.

[6] F. Alrowaie, R. B. Gopaluni, and K. E. Kwok, “Alarm design for nonlinear stochastic systems,” in *Proc. 11th World Congr. Intell. Control Autom.*, Shenyang, China, Jun. 29–Jul. 4 2014, 2014, pp. 472–479.

[7] API, 2010, *API-1167: Pipeline SCADA Alarm Management,* American Petroleum Institute. Washington, DC, USA.

[8] B. M. S. Arifin and M. A. A. S. Choudhury, “An alternative approach of risk analysis for multivariable alarm system,” *J. Chem. Eng.*, vol. 26, pp. 75–79, 2011.

[9] R. K. Arjomandi and K. Salahshoor, “Development of an efficient alarm management package for an industrial process plant,” in *Proc. Chinese Control Decision Conf.*, Mianyang, China, May 23–25, 2011, pp. 1875–1880, (2011).

[10] ASM, 2009, *ASM Consortium Guideline: Effective Alarm Management Practice,* Abnormal Situations Management Consortium.

[11] E. Barocio, B. C. Pal, D. Fabozzi, and N. F. Thornhill, “Detection and visualization of power system disturbances using principal component analysis,” in *Proc. IREP Symp.-Bulk Power Syst. Dyn. Control CIX (IREP)*, Rethymnon, Greece, Aug. 25–30, 2013, pp. 1–10.

[12] C. Basu, K. Das, J. Hazra, and D. P. Seetharam, “Enhancing wide-area monitoring and control with intelligent alarm handling,” in *Proc. 4th IEEE PES Eur. Innovative Smart Grid Technol.*, Copenhagen, Denmark, Oct. 6–9, 2013, pp. 1–5.

[13] M. Bauer, J. W. Cox, M. H. Caveness, J. J. Downs, and N. F. Thornhill, “Finding the direction of disturbance propagation in a chemical process using transfer entropy,” *IEEE Trans. Control Syst. Technol.*, vol. 15, no. 1, pp. 12–21, Jan. 2007.

[14] M. Bauer and N. F. Thornhill, “A practical method for identifying the propagation path of plant-wide disturbances,” *J. Process Control*, vol. 18, pp. 707–719, 2008.

[15] D. Beebe, S. Ferrer, and D. Logerot, “The connection of peak alarm rates to plant incidents and what you can do to minimize,” *Process Safety Progress*, vol. 32, pp. 72–77, 2013.

[16] E. Blaauwgeers, L. Dubois, and L. Ryckaert, “Real-time risk estimation for better situational awareness,” in *Proc. 12th IFAC Symp. Anal., Des., Eval. Human-Mach. Syst. V*, Las Vegas, NV, USA, Aug. 11–15, 2013, pp. 232–239.

[17] M. L. Bransby and J. Jenkinson, *The Management of Alarm Systems,* Health and Safety Executive, 1998.

[18] E. H. Bristol, “Improved process control alarm operation,” *ISA Trans.*, vol. 40, pp. 191–205, 2001.

[19] R. Brooks, R. Thorpe, and J. Wilson, “A new method for defining and managing process alarms and for correcting process operation when an alarm occurs,” *J. Hazardous Mater.*, vol. 115, pp. 169–174, 2004.

[20] D. C. C. Brown, “Alarm management experience in BP oil,” *IEE Colloquium on Best Practices in Alarm Management*, Mar. 25, 1998.

[21] N. Brown, “Alarm management/The EEMUA guidelines in practice,” *Meas. Control*, vol. 36, pp. 114–119, 2003.

[22] E. Burnell and C. R. Dicken, “Handling of repeating alarms,” *IEE Colloquium on Stemming the Alarm Flood*. London, U.K., 1997.

[23] S. Cauvin, M. O. Cordier, C. Dousson, P. Laborie, F. Levy, J. Montmain, M. Porcheron, I. Servet, and L. Trave-Massuyes, “Monitoring and alarm interpretation in industrial environments,” *AI Commun.*, vol. 11, pp. 139–173, 1998.

[24] I. M. Cecilioa, J. R. Ottewillb, J. Pretlovec, and N. F. Thornhill, “Nearest neighbors method for detecting transient disturbances in process and electromechanical systems,” *J. Process Control*, vol. 24, pp. 1382–1393, 2014.

[25] Y. Chang, F. Khana, and S. Ahmed, “A risk-based approach to design warning system for processing facilities,” *Process Safety Environ. Protection*, vol. 89, pp. 310–316, 2011.

[26] S. Charbonnier, C. Garcia-Beltan, Cadet, Catherine, and S. Gentil, “Trends extraction and analysis for complex system monitoring and decision support,” *Eng. Appl. Artif. Intell.*, vol. 18, pp. 21–36, 2005.

[27] S. Charbonnier and F. Portet, “A self-tuning adaptive trend extraction method for process monitoring and diagnosis,” *J. Process Control*, vol. 22, pp. 1127–1138, 2012.

[28] S. Charbonnier, N. Bouchair, and P. Gayet, “Fault isolation by comparing alarm lists using a symbolic sequence matching algorithm,” in *Proc. 19th IFAC World Congr.*, Cape Town, South Africa, Aug. 24–29, 2014, pp. 7085–7090.

[29] Y. Cheng, I. Izadi, and T. Chen, “Optimal alarm signal processing: Filter design and performance analysis,” *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 2, pp. 446–451, Apr. 2013a.

[30] Y. Cheng, I. Izadi, and T. Chen, “Pattern matching of alarm flood sequences by a modified Smith-Waterman algorithm,” *Chem. Eng. Res. Des.*, vol. 91, pp. 1085–1094, 2013b.

[31] S. W. Cheon, S. H. Chang, and H. Y. Chung, “Development strategies of an expert system for multiple alarm processing and diagnosis in nuclear power plants,” *IEEE Trans. Nuclear Sci.*, vol. 40, no. 1, pp. 21–30, Feb. 1993.

[32] L. H. Chiang and R. D. Braatz, “Process monitoring using causal map and multivariate statistics: Fault detection and identification,” *Chemometrics Intell. Lab. Syst.*, vol. 65, pp. 159–178, 2003.

[33] R. Dahlgren, G. Rosenwald, C. C. Liu, S. Muchlinski, A. Eide, and D. Sobajic, “Model-based synthesis and suppression of transformer alarms in a control center environment,” *IEEE Trans. Power Delivery*, vol. 13, no. 3, pp. 843–848, Jul. 1998.

[34] P. Dalapatu, S. Ahmed, and F. Khan, “Alarm allocation for event-based process alarm systems,” in *Proc. 10th IFAC Int. Symp. Dyn. Control of Process Syst.*, Mumbai, India, Dec. 18–20, 2013, pp. 815–820, (2013).

[35] F. Dashlstrand, “Consequence analysis theory for alarm analysis,” *Knowledge-Based Syst.*, vol. 15, pp. 27–36, 2002.

[36] G. J. Di Geronimo Gil, D. B. Alabi, O. E. Iyun, and N. F. Thornhill, “Merging process models and plant topology,” in *Proc. Adv. Control Ind. Process. (ADCONIP'11)*, Hangzhou, China, May 23–26, 2011.

[37] L. Dubois, J. M. Fort, P. Mack, and L. Ryckaert, “Advanced logic for alarm and event processing: Methods to reduce cognitive load for control room operators,” in *Proc.11th IFAC Symp. Anal., Des., Eval. Human-Mach. Syst.*, Valenciennes, France, Aug. 31–Sep. 3 2010, (2010).

[38] P. Duan, F. Yang, T. Chen, and S. L. Shah, “Direct causality detection via the transfer entropy approach,” *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 6, pp. 2052–2066, Nov. 2013.

[39] P. Duan, T. Chen, S. L. Shah, and F. Yang, “Methods for root cause diagnosis of plant-wide oscillations,” *AIChE Journal*, vol. 60, pp. 2019–2034, 2014.

[40] P. Duan, T. Chen, S. L. Shah, and F. Yang, “Transfer zero-entropy and its application for capturing cause and effect relationship between variables,” *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 3, pp. 855–867, May 2015.

[41] EEMUA, 2013, *EEMUA-191: Alarm Systems—A Guide to Design, Management and Procurement,* Engineering Equipment and Materials Users Association. London, U.K.

[42] EPRI, *EPRI-1010076: Advanced Control Room Alarm System: Requirements and Implementation Guidance,* Electric Power Research Institute. Charlotte, NC, USA, 2005.

[43] J. Folmer, D. Pantforder, and B. Vogel-Heuser, “An analytical alarm flood reduction to reduce operator's workload,” *Human-Comput. Interactions, Part IV, HCII*, vol. 6764, LNCS, pp. 297–306, 2011.

[44] J. Folmer and B. Vogel-Heuser, “Computing dependent industrial alarms for alarm flood reduction,” in *Proc. 9th Int. Multi-Conf. Syst., Signals Devices*, Mar. 20–23, 2012, pp. 1–6.

[45] J. Folmer, F. Schuricht, and B. Vogel-Heuser, “Detection of temporal dependencies in alarm time series of industrial plants,” in *Proc. 19th IFAC World Congr.*, Cape Town, South Africa, Aug. 24–29, 2014, pp. 1802–1807.

[46] Z. Geng, Q. Zhu, and X. Gu, “A fuzzy clustering-ranking algorithm and its application for alarm operating optimization in chemical processing,” *Process Safety Progress*, vol. 24, pp. 66–75, 2005.

[47] S. Gigi and A. K. Tangirala, “Quantitative analysis of directional strengths in jointly stationary linear multivariate processes,” *Bio. Cybern.*, vol. 103, pp. 119–133, 2010.

[48] W. Guo, F. Wen, Z. Liao, L. Wei, and J. Xin, “An analytic model-based approach for power system alarm processing employing temporal constraint network,” *IEEE Trans. Power Delivery*, vol. 25, no. 4, pp. 2435–2447, Oct. 2010.

[49] A. Gupta, A. Giridhar, V. Venkatasubramanian, and G. V. Reklaitis, “Intelligent alarm management applied o continuous pharmaceutical tablet manufacturing: An integrated approach,” *Ind. Eng. Chem. Res.*, vol. 52, pp. 12357–12368, 2013.

[50] A. Henningsen and J. P. Kemmerer, “Intelligent alarm handling in cement plants,” *IEEE Ind. Appl. Mag.*, vol. 1, no. 5, pp. 9–15, Sep./Oct. 1995.

[51] M. Hollender and C. Beuthel, “Intelligent alarming,” *ABB Review*, pp. 20–23, Feb. 2007.

[52] B. Hollifield and E. Habibi*, The Alarm Management Handbook*, 2nd ed. Houston, TX, USA: PAS, 2010.

[53] HSE, 1997, *The Explosion and Fires at the Texaco Refinery, Milford Haven, 24 July 1994: A Report of the Investigation by the Health and Safety Executive into the Explosion and Fires on the Pembroke Cracking Company Plant at the Texaco Refinery, Milford Haven on 24 July 1994,* Health and Safety Executive.

[54] HSE, 2008, *The Final Report of the Major Incident Investigation Board,* Health and Safety Executive.

[55] A. J. Hugo, “Estimation of alarm deadbands,” in *Proc. 7th IFAC Symp. Fault Detection, Supevision. Safety Tech. Process.*, Barcelona, Spain, Jun. 30–Jul. 3 2009.

[56] S. L. Hwang, J. T. Lin, G. F. Liang, Y. J. Yau, T. C. Yenn, and C. C. Hsu, “Application control chart concepts of designing a pre-alarm system in the nuclear power plant control room,” *Nuclear Eng. Des.*, vol. 238, pp. 3522–3527, 2008.

[57] IEC, IEC-68682: Management of Alarm Systems for the Process Industries 2014, International Electrotechnical Commission.

[58] ISA, 2009, *ANSI/ISA-18.2: Management of Alarm Systems for the Process Industries.* International Society of Automation. Durham, NC, USA.

[59] Isermann, *Fault-Diagnosis Applications: Model-Based Condition Monitoring*. New York, NY, USA, Springer, 2011.

[60] I. Izadi, S. L. Shah, D. Shook, and T. Chen, “An introduction to alarm analysis and design,” in *Proc. 7th IFAC Symp. Fault Detection, Supervision Safety Tech. Process.*, Barcelona, Spain, Jun. 30–Jul. 3 2009, pp. 645–650.

[61] G. Jang, S. Suh, S. Kim, Y. Suh, and J. Park, “A proactive alarm reduction method and its human factors validation test for a main control room for SMART,” *Ann. Nuclear Energy*, vol. 51, pp. 125–134, 2013.

[62] E. Jerhotova, M. Sikora, and P. Stluka, “Dynamic alarm management in next generation process control systems,” in *APMS 2012, Part II, IFIP AICT 398*, 2013, pp. 224–231.

[63] H. Jiang, R. Patwardhan, and S. L. Shah, “Root cause diagnosis of plant-wide oscillations using the concept of adjacency matrix,” *J. Process Control*, vol. 19, pp. 1347–1354, 2009.

[64] M. Kezunovic and Y. Guan, “Intelligent alarm processing: From data intensive to information rich,” in *Proc. 42nd Hawaii Int. Conf. Syst. Sci.*, HI, USA, Jan. 5–8, 2009, pp. 1–8, (2009).

[65] I. S. Kim, “Computerized systems for online management of failures: A state-of-the-art discussion of alarm systems and diagnostic systems applied in the nuclear industry,” *Reliability Eng. Syst. Safety*, vol. 44, pp. 279–295, 1994.

[66] D. S. Kirschen and B. F. Wollenberg, “Intelligent alarm processing in power systems,” *Proc. IEEE*, vol. 80, pp. 663–672, 1992.

[67] S. R. Kondaveeti, S. L. Shah, and I. Izadi, “Application of multivariate statistics for efficient alarm generation,” in *Proc. 7th IFAC Symp. Fault Detection, Supervision Safety Tech. Process.*, Barcelona, Spain, Jun. 30–Jul. 3 2009, pp. 657–662.

[68] S. R. Kondaveeti, I. Izadi, S. L. Shah, D. Shook, R. Kadali, and T. Chen, "Quantification of alarm chatter based on run length distributions," *Chem. Eng. Res. Des.*, vol. 91, pp. 2550–2558, 2013.

[69] S. R. Kondaveeti, I. Izadi, S. L. Shah, T. Black, and T. Chen, "Graphical tools for routine assessment of industrial alarm systems," *Comput. Chem. Eng.*, vol. 46, pp. 39–47, 2012.

[70] S. Kordic, C. P. Lam, J. Xiao, and H. Li, "Using event-intervals to analyze alarm sequences in a chemical plant," *J. Frontiers Comput. Sci. Technol.*, vol. 3, pp. 267–281, 2009.

[71] J. C. Laberge, P. Bullermer, M. Tolsma, and D. V. C. Reising, "Addressing alarm flood situations in the process industries through alarm summary display design and alarm response strategy," *Int. J. Ind. Ergon.*, vol. 44, pp. 395–406, 2014.

[72] R. Landman, J. Kortela, Q. Sun, and S.-L. Jämsä-Jounela, "Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity," *Comput. Chem. Eng.*, vol. 71, pp. 446–456, 2014.

[73] J. E. Larsson, B. Ohman, A. Calzada, C. Nihlwing, H. Jokstad, L. I. Kristianssen, J. Kvalem, and M. Lind, "A revival of the alarm system: Making the alarm list useful during incidents," in *Proc. 5th Int. Topical Meeting on Nuclear Plant Instrumentation Controls, and Human Machine Interface Technol.*, Albuquerque, NM, USA, Nov. 12–16, 2006, pp. 711–716.

[74] J. W. Lee, J. T. Kim, J. C. Park, I. K. Hwang, and S. P. Lyu, "Computer-based alarm processing and presentation methods in nuclear power plants," *World Acad. Sci., Eng. Technol.*, vol. 41, pp. 594–598, 2010.

[75] J. Liu, K. W. Lim, W. K. Ho, K. C. Tan, R. Srinivasan, and A. Tay, "The intelligent alarm management system," *IEEE Softw.*, vol. 20, pp. 66–71, 2003.

[76] X. Liu, M. Noda, and H. Nishitani, "Evaluation of plant alarm systems by behavior simulation using a virtual subject," *Comput. Chem. Eng.*, vol. 34, pp. 374–386, 2010.

[77] D. Macdonald, *Practical Hazops, Trips and Alarms*. New York, NY, USA: Elsevier, 2004.

[78] M. R. Maurya, R. Rengaswamy, and V. Venkatasubramanian, "A systematic framework for the development and analysis of signed digraphs for chemical processes. 1. Algorithms and analysis," *Ind. Eng. Chem. Res.*, vol. 42, pp. 4789–4810, 2003.

[79] M. R. Maurya, R. Rengaswamy, and V. Venkatasubramanian, "A signed directed graph and qualitative trend analysis-based framework for incipient fault diagnosis," *Chem. Eng. Res. Des.*, vol. 85, pp. 1407–1422, 2007.

[80] J. R. McDonald, G. M. Burt, and D. J. Young, "Alarm processing and fault diagnosis using knowledge based systems for transmission and distribution network control," *IEEE Trans. Power Syst.*, vol. 7, no. 3, pp. 1292–1298, Aug. 1992.

[81] H. Miao, M. Sforna, and C. C. Liu, "A new logic-based alarm analyzer for online operational environment," *IEEE Trans. Power Syst.*, vol. 11, no. 3, pp. 1600–1606, Aug. 1996.

[82] E. Naghoosi, I. Izadi, and T. Chen, "Estimation of alarm chattering," *J. Process Control*, vol. 21, pp. 1243–1249, 2011.

[83] NAMUR, *NAMUR-NA-102: Alarm Management,* Standardization Association for Measurement and Control in Chemical Industries, 2003.

[84] C. Nihlwing and M. Kaarstad, "The development and usability test of a state based alarm system for a nuclear power plant simulator," in *Proc. NPIC and HMIT*, San Diego, CA, USA, Jul. 22–26, 2012.

[85] I. Nimmo, "Rescue your plant from alarm overload," *Chem. Process.*, pp. 28–33, 2005.

[86] M. Noda, F. Higuchi, T. Takai, and H. Nishitani, "Event correlation analysis for alarm system rationalization," *Asia-Pac. J. Chem. Eng.*, vol. 6, pp. 497–502, 2011.

[87] NRC, *NUREG/CR-61056684: Advanced Alarm Systems: Revision of Guidance and Its Technical Basis* 2000, Nuclear Regulatory Commission. Washington, DC, USA.

[88] A. Pariyani, W. D. Seider, U. G. Oktem, and M. Soroush, "Incidents investigation and dynamic analysis of large alarm databases in chemical plants: A fluidized-catalytic-cracking unit case study," *Ind. Eng. Chem. Res.*, vol. 49, pp. 8062–8079, 2010.

[89] D. Rothenberg, *Alarm Management for Process Control*. New York, NY, USA: Momentum Press, 2009.

[90] M. Schleburg, L. Christiansen, N. F. Thornhill, and A. Fay, "A combined analysis of plant connectivity and alarm logs to reduce the number of alerts in an automation system," *J. Process Control*, vol. 23, pp. 839–851, 2013.

[91] Z. Simeu-Abazi, A. Lefebvre, and J. P. Derain, "A methodology of alarm filtering using dynamic fault tree," *Reliability Eng. Syst. Safety*, vol. 96, pp. 257–266, 2011.

[92] J. C. S. Souza, E. M. Meza, M. T. Schilling, and M. B. D. C. Filho, "Alarm processing in electrical power systems through a neuro-fuzzy approach," *IEEE Trans. Power Delivery*, vol. 19, no. 2, pp. 537–544, Apr. 2004.

[93] R. Srinivasan, J. Liu, K. W. Lim, K. C. Tan, and W. K. Ho, "Intelligent alarm management in a petroleum refinery," *Hydrocarbon Process.*, vol. 11, pp. 47–52, 2004.

[94] T. Stauffer and P. Clarke, "Using alarms as a layer of protection," in *Proc. 8th Global Congr. Process Safety*, Houston, TX, USA, Apr. 1–4, 2012.

[95] K. Takeda, T. Hamaguchi, M. Noda, N. Kimura, and T. Itoh, "Use of two-layer cause-effect model to select source of signal in plant alarm system," in *KES2010, Part II*, R. Setchi et al., Ed., 2010, vol. 6277, LNAI, pp. 381–388.

[96] K. Takeda, T. Hamaguchi, N. Kimura, and M. Noda, "A method of designing plant alarm system based on first alarm alternative signals for each assumed plant malfunction," in *Proc. 6th Int. Conf. Process Syst. Eng. (PSE ASIA)*, Kuala Lumpur, Malaysia, Jun. 25–27, 2013.

[97] A. Tchamova and J. Dezert, "Intelligent alarm classification based on DSmT," in *Proc. 6th IEEE Int. Conf. Intell. Syst.*, Sofia, Bulgaria, Sep. 6–8, 2012, pp. 120–125.

[98] J. Thambirajah, L. Benabbas, M. Bauer, and N. F. Thornhill, "Cause-and-effect analysis in chemical processes utilizing XML, plant connectivity and quantitative process history," *Comput. Chem. Eng.*, vol. 33, pp. 503–512, 2009.

[99] N. F. Thornhill, J. W. Cox, and M. A. Paulonis, "Diagnosis of plant-wide oscillation through data-driven analysis and process understanding," *Control Eng. Practice*, vol. 11, pp. 1491–1490, 2003.

[100] N. F. Thornhill, "Finding the source of nonlinearity in a process with plant-wide oscillation," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 3, pp. 434–443, May 2005.

[101] C. Timms, "A methodology for alarm classification and prioritization," in *Proc. Int. Conf. Human Interfaces in Control Rooms, Cockpits and Command Centers*, Bath, U.K., Jun. 21–23, 1999, pp. 222–228.

[102] C. Timms, "Hazards equal trips or alarms or both," *Process Safety Environ. Protection*, vol. 87, pp. 3–13, 2009.

[103] U. S. Tolga, J. Niels, L. Morten, and J. S. Bay, "Fundamental principles of alarm design," *Nuclear Safety and Simulations*, vol. 2, pp. 44–51, 2011.

[104] T. Varga, F. Szeifert, and J. Abonyi, "Detection of safe operating regions: a novel dynamic process simulator based predictive alarm management approach," *Ind. Eng. Chem. Res.*, vol. 49, pp. 658–668, 2010.

[105] J. Wang and T. Chen, "An online method for detection and reduction of chattering alarms due to oscillation," *Comput. Chem. Eng.*, vol. 54, pp. 140–150, 2013.

[106] J. Wang and T. Chen, "An online method to remove chattering and repeating alarms based on alarm durations and intervals," *Comput. Chem. Eng.*, vol. 67, pp. 43–52, 2014.

[107] J. Wang, J. Xu, and D. Zhu, "Online root-cause analysis of alarms in discrete Bayesian networks with known structures," in *Proc. 11th World Congr. Intell. Control Autom.*, Shenyang, China, Jun. 29–Jul. 4 2014, pp. 467–472.

[108] J. Wang, H. Li, J. Huang, and C. Su, "A data similarity based analysis to consequential alarms of industrial processes," *J. Loss Prevention in the Process Industries*, vol. 35, pp. 29–34, 2015.

[109] F. Wen and C. S. Chang, "Tabu search approach to alarm processing in power systems," *IEE Proc.-Gener. Transm. Distrib.*, vol. 144, pp. 31–38, 1997.

[110] F. Wen, C. S. Chang, and W. Fu, "New approach to alarm processing in power systems based upon the set covering theory and a refined genetic algorithm," *J. Electric Mach. Power Syst.*, vol. 26, pp. 53–67, 1998.

[111] L. Wei, W. Guo, F. Wen, G. Ledwich, Z. Liao, and J. Xin, "An online intelligent alarm-processing system for digital substation," *IEEE Trans. Power Delivery*, vol. 26, no. 3, pp. 1615–1624, Jul. 2011.

[112] J. Xu, J. Wang, I. Izadi, and T. Chen, "Performance assessment and design for univariate alarm systems based on FAR, MAR and AAD," *IEEE Trans. Autom. Sci. Eng.*, vol. 9, no. 2, pp. 296–307, Apr. 2012.

[113] S. Xu, A. Adhitya, and R. Srinivasan, "Hybrid mode-based framework for alarm anticipation," *Ind. Eng. Chem. Res.*, vol. 53, pp. 5182–5193, 2014.

[114] F. Yamanaka and T. Nishiya, "Application of the intelligent alarm system for the plant operation," *Comput. Chem. Eng.*, vol. 21, pp. 625–630, 1997.

[115] L. Yan, X. Liu, M. Noda, and H. Nishitani, "Systematic design approach for plant alarm systems," *J. Chem. Eng. Jpn.*, vol. 40, pp. 765–772, 2007.

[116] F. Yang, D. Xiao, and S. L. Shah, "Optimal sensor location design for reliable fault detection in presence of false alarms," *Sensor*, vol. 9, pp. 8579–8592, 2009.

[117] F. Yang, S. L. Shah, and D. Xiao, "Correlation analysis of alarm data and alarm limit design for industrial processes," in *Proc. Amer. Control Conf.*, Baltimore, MD, USA, Jun. 30–Jul. 2 2010, pp. 5850–5855.
[118] F. Yang, S. L. Shah, D. Xiao, and T. Chen, "Improved correlation analysis and visualization of industrial alarm data," *ISA Trans.*, vol. 51, pp. 499–506, 2012.
[119] F. Yang, S. L. Shah, and D. Xiao, "Signed directed graph based modeling and its validation from process knowledge and process data," *Int. J. Appl. Math. Comput. Sci.*, vol. 22, pp. 387–392, 2012.
[120] Z. Yang, J. Wang, and T. Chen, "Detection of correlated alarms based on similarity coefficients of binary data," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 4, pp. 1014–1025, Oct. 2013.
[121] F. Yang, P. Duan, S. L. Shah, and T. Chen, *Capturing Connectivity and Causality in Complex Industrial Processes*. New York, NY, USA: Springer, 2014.
[122] S. Y. Yim, H. G. Ananthakumar, L. Benabbas, A. Horch, R. Drath, and N. F. Thornhill, "Using process topology in plant-wide control loop performance assessment," *Comput. Chem. Eng.*, vol. 31, pp. 86–99, 2006.
[123] D. J. Young, K. L. Lo, J. R. McDonald, R. Howard, and J. Rye, "Development of a practical expert system for alarm processing," *IEE Proc.-C*, vol. 139, pp. 437–447, 1992.
[124] T. Yuan and S. J. Qin, "Root cause diagnosis of plant-wide oscillations using Granger causality," *J. Process Control*, vol. 24, pp. 450–459, 2014.
[125] Y. Yuki, "Alarm system optimization for increasing operations productivity," *ISA Trans.*, vol. 41, pp. 383–387, 2002.
[126] H. Zang and H. Li, "Optimization of process alarm thresholds: A multidimensional kernel density estimation approach," *Process Safety Progress*, vol. 33, pp. 292–298, 2014.
[127] Q. Zhu and Z. Geng, "A new fuzzy clustering-ranking algorithm and its application in process alarm management," *Chinese J. Chem. Eng.*, vol. 13, pp. 477–483, 2005.
[128] J. Zhu, Y. Shu, J. Zhao, and F. Yang, "A dynamic alarm management strategy for chemical process transitions," *J. Loss Prevention in the Process Industries*, vol. 30, pp. 207–218, 2014.
[129] J. Zhu, J. Zhao, and F. Yang, "Dynamic risk analysis with alarm data to improve process safety using Bayesian network," in *Proc. 11th World Congr. Intell. Control Autom.*, Shenyang, China, Jun. 29–Jul. 4 2014, pp. 461–466.

Jiandong Wang received the B.E. degree in automatic control from the Beijing University of Chemical Technology, Beijing, China, in 1997, and the M.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Alberta, Edmonton, AB, Canada, in 2003 and 2007, respectively.

He is presently a Professor with the Department of Industrial Engineering and Management, College of Engineering, Peking University, Beijing, China. From 1997 to 2001, he was a Control Engineer with the Beijing Tsinghua Energy Simulation Company, Beijing. From February 2006 to August 2006, he was a Visiting Scholar at the Department of System Design Engineering, Keio University, Japan. His research interests include system identification, industrial alarm systems, control loop performance monitoring and optimization, and their applications to industrial processes.

Dr. Wang served as an Associate Editor for *Systems and Control Letters*.

Fan Yang (M'06) received the B.Eng. degree in automation and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2002 and 2008, respectively.

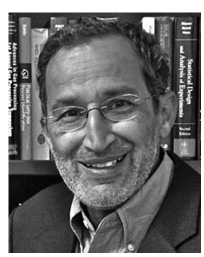
He joined the Department of Automation at Tsinghua University, in 2011, after working as a Post-Doctoral Fellow with Tsinghua University and the University of Alberta, Edmonton, AB, Canada. He is currently an Associate Professor with Tsinghua University. He has coauthored a brief entitled "Capturing Connectivity and Causality in Complex Industrial Processes" (Springer). His current research interests include topology modeling of large-scale processes, abnormal events monitoring, process hazard analysis, and smart alarm management.

Dr. Yang was a recipient of the Young Research Paper Award from the IEEE Control Systems Society Beijing Chapter in 2006, the Outstanding Graduate Award from Tsinghua University in 2008, and the Teaching Achievement Awards from Tsinghua University in 2012 and 2014.

Tongwen Chen (F'06) He received the B.Eng. degree in automation and instrumentation from Tsinghua University (Beijing) in 1984, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Toronto in 1988 and 1991, respectively.

He is presently a Professor of Electrical and Computer Engineering at the University of Alberta, Edmonton, Canada. His research interests include computer and network-based control systems, process safety and alarm systems, and their applications to the process and power industries.

Prof. Chen served as an Associate Editor for several international journals, including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *Automatica*, and *Systems and Control Letters*. He is a Fellow IFAC, as well as the Canadian Academy of Engineering.



Sirish L. Shah (M'76) is on faculty of the University of Alberta, Edmonton, Canada, where he held the NSERC-Matrikon-Suncor-iCORE Senior Industrial Research Chair in Computer Process Control from 2000 to 2012. He has held visiting appointments at Oxford University and Balliol College as a SERC Fellow, Kumamoto University (Japan) as a Senior Research Fellow of the Japan Society for the Promotion of Science (JSPS), the University of Newcastle, Australia, IIT-Madras, India, and the National University of Singapore. He has coauthored three books, the first titled "Performance Assessment of Control Loops: Theory and Applications," a second book titled book titled "Diagnosis of Process Nonlinearities and Valve Stiction: Data Driven Approaches," and a more recent brief monograph titled "Capturing Connectivity and Causality in Complex Industrial Processes." The main area of his current research is process and performance monitoring, system identification and design, analysis and rationalization of alarm systems.