Timely alarms

Abstract

IEC 62682 introduced the word 'timely' in the definition of an alarm. This paper discusses what is meant, how timeliness can be measured and how alarm systems can be improved to meet this criterium.

Introduction

The definition of an alarm was set in ANSI/ISA 18.2 2009 as follows:

An audible and/or visible means of indicating to the operator an equipment malfunction, process deviation, or abnormal condition requiring a response.

This definition was based on the recommendations of EEMUA 191 and the "Abnormal Situation Management – Effective Automation to Improve Operator Performance" of the ASM® consortium.

Progressing insight and field experience of those who contributed to the preparation of the IEC 62682 standard have led to the improvement of the above definition to:

An audible and/or visible means of indicating to the operator an equipment malfunction, process deviation, or abnormal condition requiring a timely response.

Timely, however, is nowhere defined in the standards (IEC 62682 and ISA 18.2). EEMUA 191 specifies timely as: "not long before any response is needed or not too late to do anything". Timely is defined in the Cambridge dictionary as "happening at the best possible moment".

It is left to the individual sites to define timely in their alarm philosophy. So, what is the best possible moment? This paper discusses different aspects of this issue.

Alarm lifecycle

An alarm exists when it is annunciated to an operator display. It requires the attention of the operator to take some action. This is important: the alarm does not serve to only receive attention of the operator, the operator is required to perform an action. The operator acknowledges he has read and understood the alarm and takes responsibility to undertake some actions to make the alarm disappear. This is also important: acknowledging there is an alarm is not the required action; actually, operators should be conscient that acknowledging means taking responsibility! When these actions are successful, the alarm disappears or 'clears'. This provides the following timeline:

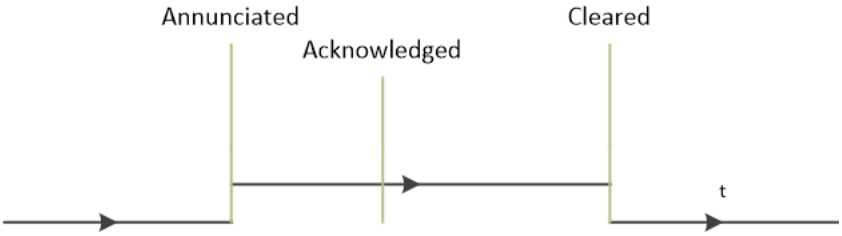

Figure 1: alarm timeline

A proper alarm and event reporting and assessment tool will collect these events and store them in a database for further analysis.

Measuring the timeliness of alarms

Can the monitoring and assessment tools of a site assist in defining the timeliness of alarms in the current alarm system? The monitoring and assessment tool used, should assist in the elimination of alarms that are too early or too late.

Fleeting alarms

Fleeting alarms, alarms that pop up for just a short while and then disappear without any operator action, give the operator the impression he is too late. Unlike chattering alarms, which appear and disappear frequently and are perceived as nuisance, fleeting alarms could distract the operator's attention. To avoid such distraction and resulting human errors, these should be removed using techniques such as persistence delay time.

The ProcessVue Analyser software from MAC Solutions (UK) Ltd can provide a report detailing the amount of time spent in an alarm state and plots this on a chart, example below:

Tag name	OFFNRM	Priority 1
Duration	Count	Percentage
< 1 Seconds	0	0%
< 2 Seconds	4	5.88%
< 3 Seconds	12	17.65%
< 4 Seconds	15	22.06%
< 5 Seconds	22	32.35%
< 6 Seconds	24	35.29%
< 7 Seconds	24	35.29%
< 8 Seconds	24	35.29%
< 9 Seconds	26	38.24%
< 10 Seconds	31	45.59%
< 11 Seconds	36	52.94%
< 12 Seconds	42	61.76%
> 12 Seconds	26	38.24%

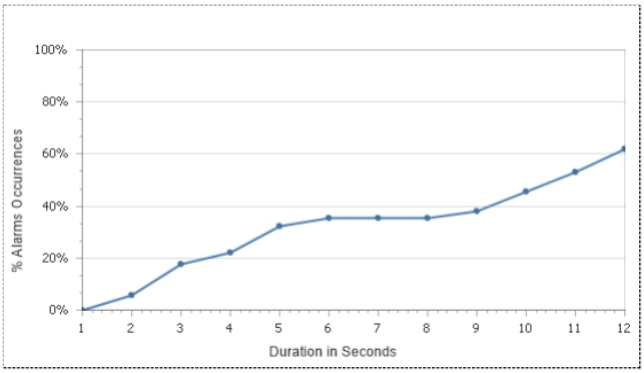

Figure 2: Example of measuring the lifetime of an alarm, courtesy of ProcessVue Analyser

Here we can see on this particular high priority alarm, 62% of alarms last less than 12 seconds. Ask the question - What actions can an operator undertake in (less than) 12 seconds?

Fleeting alarms always come too late and should undergo maintenance.

Long-standing alarms

Long standing alarms are alarms that do not go away or only after a very long time. When acknowledged by the operator, it signifies that whatever action the operator has undertaken, the alarm does not disappear. It cannot be concluded that such alarm came too early, it should be investigated whether such an alarm is really an alarm at all, as apparently in most cases the process can operate without this alarm being cleared.

Trip alarms

Another common mistake is to put alarms on every consequence threshold. When the alarm system is properly configured, the consequence threshold implies that some safety function kicks in, meaning the operator cannot perform any action to avoid the consequence. However, in many situations, it is required he performs actions to release the safety function or restart the process. That is, when the process or installation has fully returned to the safe state. So, the action to perform should only be initiated when the safe state is reached. If it takes twenty minutes or more to reach the safe state, then the trip point should not be annunciated as an alarm, but only as a notification. The alarm could be set at the point in time when the operator can start performing actions to restart.

Many trip alarms come too early and should be rationalized.

Defining the alarm limit or setpoint

From the above, it is obvious that defining the alarm setpoint will be crucial in providing the operator with sufficient time.

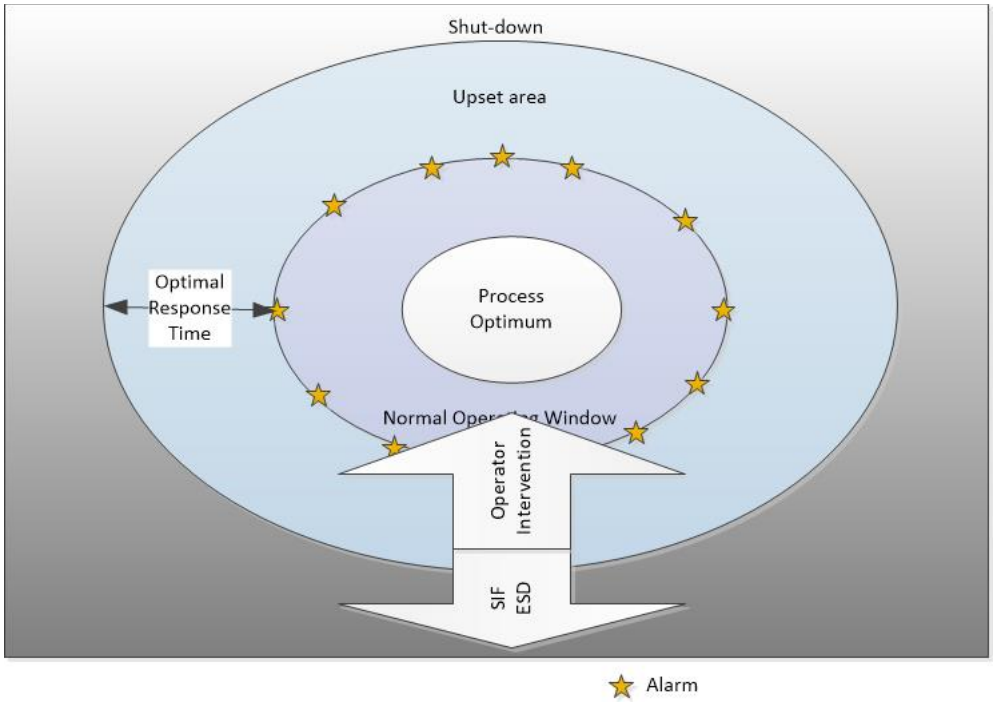

Figure 3: example of an effective alarm system according EEMUA 191 Third Edition Figure 11

What does the operator need to do?

The standards (ISA/ANSI 18.2 and IEC 62682) propose the following model:

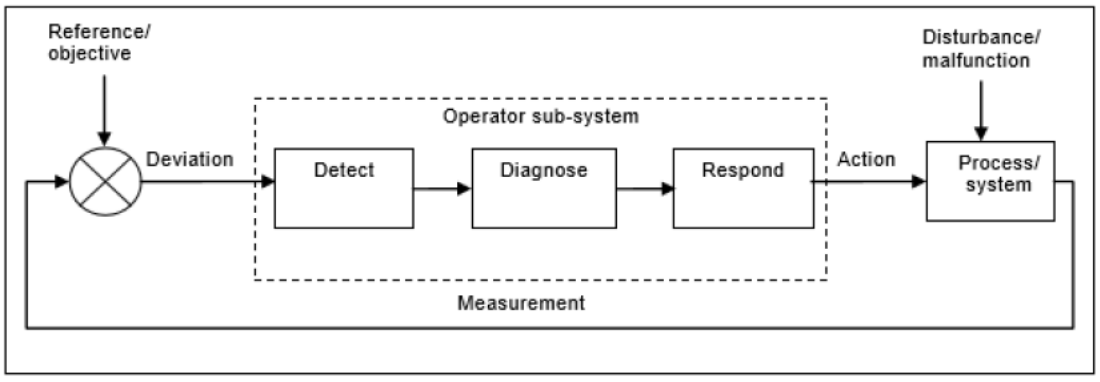

Figure 4: definition of the operator sub system in the alarm management standards (ISA 18.2, IEC 62682)

However, the operator should first be aware that a deviation requiring his attention exists. This is called situational awareness. Apart from the fact that the operator should be in front of his console, he may be distracted by other issues, like people coming into the control room and asking for permissions, managers and supervisors asking questions, telephone calls, e-mails and … other alarms.

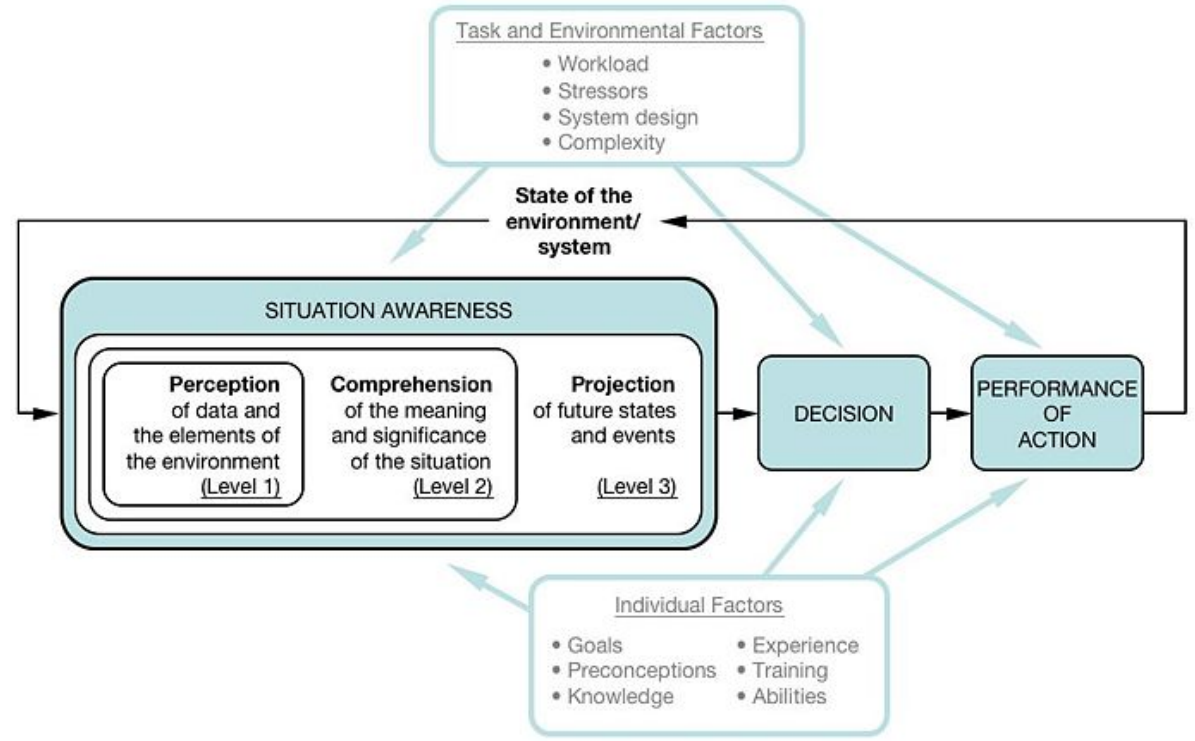

Figure 5: definition of situation awareness, courtesy by M. Endsley

Endsley[1], a Human Factors Engineering specialist, sketched the above diagram, which adds some complexity to the simple presentation in the standards. Usually, it is the control or supervisory system or the instrumentation in place, which detects the deviation and makes the operator aware of the situation by means of an alarm. Next, the operator needs to collect all required data, he needs to understand the current situation, he needs to assess what may happen depending on the actions he may take, which might include multiple actions depending on the underlying cause of the problem. He then needs to make a decision of what action(s) to take, perform the action(s) and monitor the outcome of the action(s). EEMUA 191 proposes the following figure to demonstrate the operator behavior in case of an alarm:

[1] In 1982, Endsley graduated cum laude from Texas Tech University in Lubbock, Texas with Bachelor of Science degree in Industrial Engineering.[2] In 1985, she earned a Master of Science degree in Industrial Engineering from Purdue University and earned a Doctor of Philosophy in Industrial and Systems engineering from the University of Southern California in 1990.

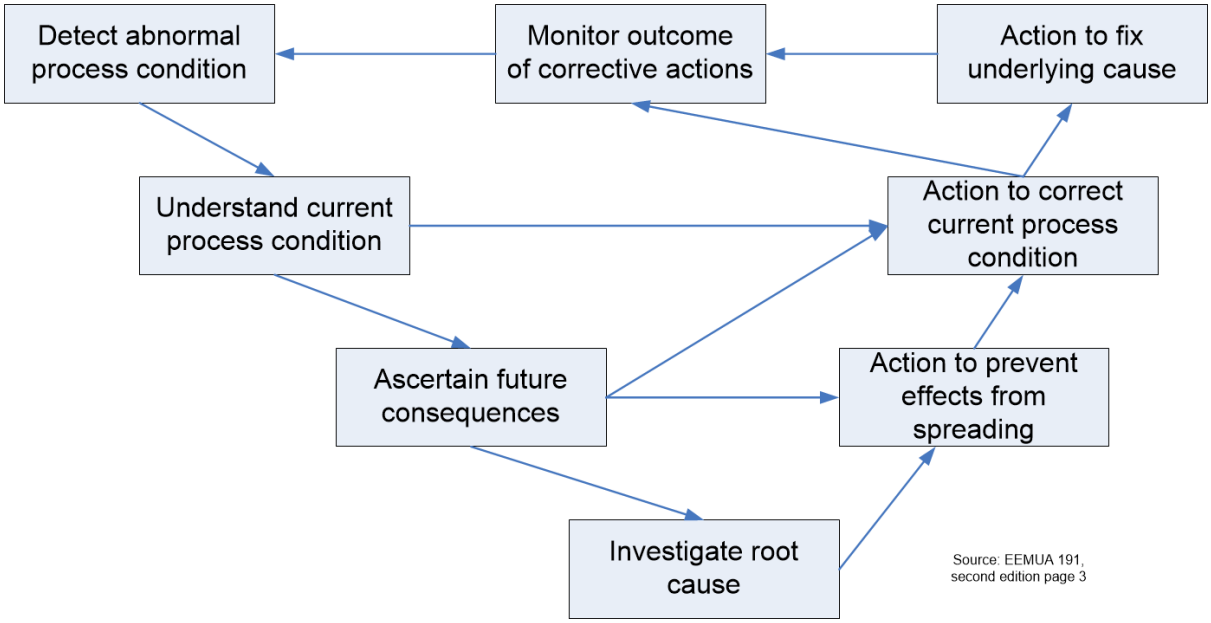

Figure 6: the role of the operator when an alarm is annunciated. Courtesy by EEMUA 191, second edition

To avoid human errors, EEMUA 191 suggests that for abnormal situations where only one action is possible, to automate the action and not generate an alarm.

To further avoid human errors (e.g., based on cognitive biases), the operators should be trained not to jump to the action part, but to first stop and think, to properly and holistically assess the situation.

Taking this into account, the rationalization team will need to assess how much time is required for the operator to acknowledge there is an alarm, evaluate the situation and take appropriate action. The next figure illustrates how the change of a (single) process variable leads to a consequence threshold, when for example no action is taken, and details each step that needs to be considered when defining the alarm threshold (alarm set point or alarm limit).

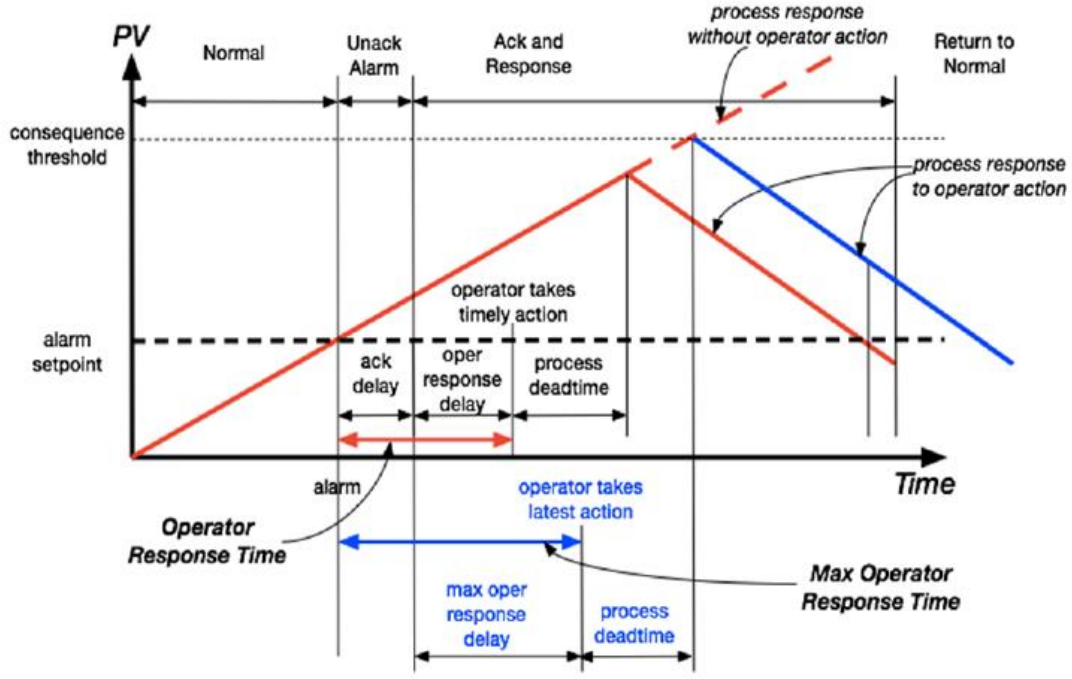

Figure 7: Operator response time, courtesy of Ian Nimmo from UCDS

Process dead time

The process dead time is the time the process needs to respond to the operator action(s). Valves do not close immediately; a heating is not stopped immediately, a rotation is not stopped immediately, and so forth. Process dead time will vary from situation to situation and can be estimated, calculated or measured.

The process deadtime is required to calculate or estimate the maximum operator response time.

Maximum operator response time

From the figure, the operator response time can be calculated from the moment the alarm is annunciated, i.e., when the process variable's value passes the alarm setpoint, to the time the process variable's value passes the consequence threshold minus the process deadtime and the acknowledgement delay, when the operator is busy with other tasks.

Required operator response time

The required operator response time cannot be derived from the figure but can be estimated based on the complexity of the alarm annunciated. How much time is required should come from the rationalization process. The actions the operator or the field engineer must undertake, how many checks need to be performed before action can be taken, the procedure to be followed when the alarm occurs, etc., are all elements that contribute to the required operator response time.

Allowable operator response time

How much time is allowed for an operator to respond, is related to the alarm philosophy. Does plant management allow the operator to go away from his console and if so, for how many minutes? Is there a (permanent) back-up available? Human beings do need to take sanatory breaks, and both factors will define this allowable away time. Today, for many plants, console operators are often alone in their control room and a back-up resource is not readily available. In such circumstances, allowable time is important!

Calculation

The sum of the allowable operator response time plus the required operator response time needs to be smaller than the maximum operator response time.

Example #1

The allowable operator response time is set to five minutes.
The process dead-time for a given alarm is one minute.
The required operator response time for this given alarm is estimated at two minutes.
The maximum operator response time for this given alarm is ten minutes:

$5 + 1 + 2 < 10$

This is fine.

Example #2

The allowable operator response time is set to five minutes.
The required operator response time for this given alarm is estimated at five minutes.
The maximum operator response time for this given alarm is ten minutes:

$5 + 5 = 10$

This alarm is prone to be too late. It should be observed how often and under what circumstances.

Example #3

The allowable operator response time is set to five minutes.
The required operator response time for this given alarm is estimated at three minutes.
The maximum operator response time for this given alarm is five minutes:

5 + 3 > 5

This alarm will most likely come too late, but it should be observed how often and under what circumstances.

Example #4

The allowable operator response time is set to five minutes.
The required operator response time for this given alarm is estimated at six minutes.
The maximum operator response time for this given alarm is five minutes:

5 + 6 > 5

This alarm is always too late, regardless the allowable time.

Process example

When a tank is filled, the input flow controller will (gradually) cut off the inflow when the level is above a given value. If the alarm setpoint is lower than the process setpoint (the level to be achieved) and higher than the level where to controller smoothly cuts the inflow, the alarm will always activate in this operating area, requiring no operator action.

During rationalization, this alarm will be classified as unnecessary because no operator action is required. Actually, it is the alarm setpoint which needs to be reconsidered. The alarm setpoint should have been defined above the cut-off limit of the control loop and enough above the process setpoint to avoid being triggered by ripple on the level surface. Consequently, it will need to be set between the cut-off level and the overflow level.

At that point, does the operator still have enough time to perform an action? Obviously, the level controller is not performing its desired task. The operator needs to check if the control loop is set to ‘manual’ or on an ill-defined simulated setpoint, he needs to check what the inflow rate is, how much time is left to overflow the tank, assess if and how he can stop the inflow, assess if he can start an outflow and consider the potential safety function in place.

Suppose he can perform all these checks in two minutes and the execution of the corresponding action can be done within a minute. The required time is therefore three minutes. Suppose the allowable ‘away’ time in the alarm philosophy is set to five minutes, then these eight minutes should be less than the maximum operator response time as sketched in Figure 7: Operator response time, courtesy of Ian Nimmo from UCDS.

In many cases, production will want to fill such a tank to its maximum capacity and safety may consider eight minutes as too risky, as the maximum operator response time in the alarm philosophy is set to ten or fifteen minutes. Consequently, an independent safety instrumented function will be implemented, enabling cutting off the inflow at a trip point. If this trip point is close the upper allowable limit, leaving the operator less than the allowable time, then there should be no alarm at all. The inflow is cut off, something that should have been accomplished by the control loop is now being performed by an independent safety function. The tank is filled to its maximum capacity and the operator should be made aware that the maximum capacity is reached and that the safety

function is or was active. This indication should not be an alarm, just a notice to the operator, which should be made sufficiently visible in the HMI to catch his attention.

Problems with moving the alarm setpoint

The figure below illustrates the problem when moving an alarm setpoint to meet the maximum operator response time as defined in the Alarm Philosophy

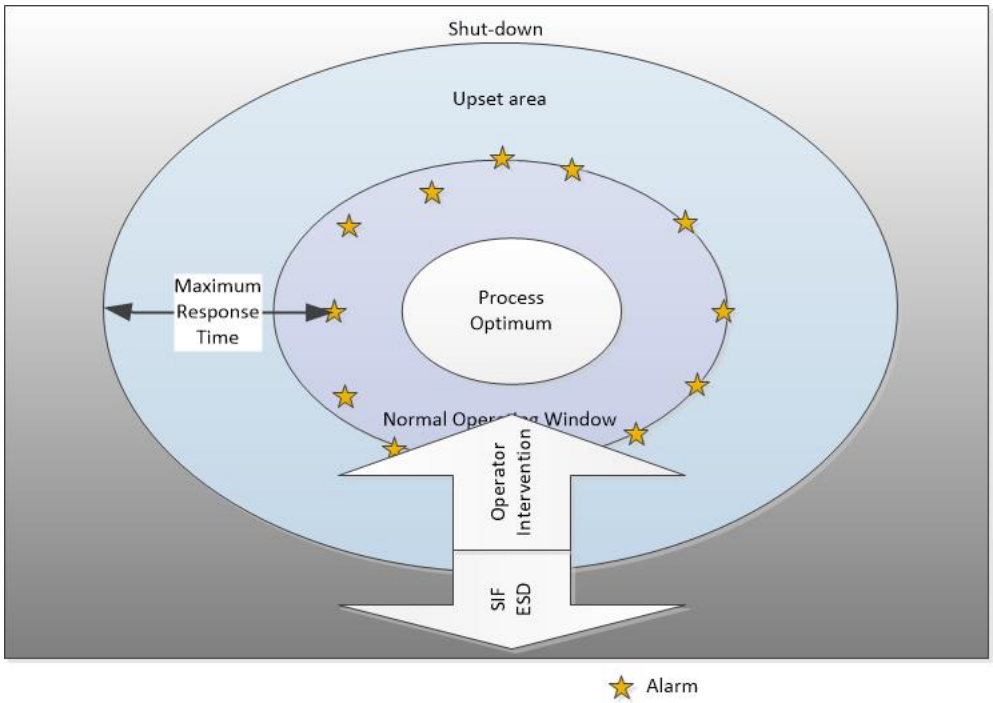

Figure 8: affective/ineffective alarm system as per Figure 11 of the Third Edition of EEMUA 191

To provide enough time for the operator to perform his actions plus the allowable time to complete other actions or tasks, a very common problem is that the alarm setpoint is moved into the normal operating window. This is usually not acceptable, as at such point there is no certainty that the measured value will move out of the normal operating window. If an alarm set point is set in the normal operating window, it will probably cause nuisance alarms. In fact, this single alarm setpoint cannot be used to trigger the desired alarm.

One solution is that during rationalization the rationalization team needs to evaluate other parameters, conditions and eventually other process values to ascertain the alarm trigger is set at the right boundary. Other parameters could be the direction of the rate of change of the process value in question, the rate of change of the process value, etc. Conditions to be evaluated could be the process condition, the configuration of the control loop (for example manually or automatic), etc. Other process values could be anything else which is measured in conjunction to the abnormal situation which is to be avoided.

Another solution is to implement a safety function which triggers the required actions without any operator intervention. This might require installing other final control elements and/or additional programming.

Advanced methods to adjust alarm trigger

If a single alarm setpoint on a single process value cannot be used as the alarm trigger, then advanced methods should be considered. ISA 18.2 TR4 provides a number of suggestions.

State based alarming

Considering process or recipe states in conjunction with how a process variable drifts or moves, is a good and recommended practice. More information can be found in section 6 of TR4.

Dynamic cause analysis

Dynamic (continuous) analysis of all potential causes of an abnormal situation can be used to trigger the alarm at the right moment, with the right information, such that operator mistakes can be avoided. More information can be found in section 7 of TR4.

Model based alarming

Use of mathematical (white) models

Mathematical (physical) (simulation) models of a process can run in parallel with the process under observation and ahead in time. Such models can predict the abnormal situation and thus provide the operator with enough time to respond.

Example: every five minutes the simulation model is launched with the current process data. It calculates the future process data for the next thirty minutes or more. When during these calculations some alarm thresholds are exceeded, a predictive alarm can be generated.

Use of artificial neural networks or black models

Data mining techniques and artificial intelligence can analyze abnormal situations stored in the process historian and provide predictions on abnormal situations and consequently provide more time for the operator to respond.

Example: the artificial neural network is fed every ten seconds with the current process data to estimate the future process data. If any of the output values exceeds an alarm setpoint, a predictive alarm can be generated.

Use of topological models and topological reasoning

In a plant different units and utilities are interconnected (think of process flows, energy flows and information flows). When an abnormal situation occurs, topological reasoning can assist in generating alarms upstream or downstream in a timely fashion, such that the operator will not be flooded with alarms he cannot do anything about. Such reasoning can also assist in suppressing consequence alarms upstream or downstream, but this aspect is subject for a future white paper.

Conclusions

Timely alarms

As EEMUA 191 specifies, a timely alarm is an alarm that does not come too late to take any action and a timely alarm is an alarm that does not come too early.

Too late means that the required time to perform all required actions is not available.

Too early means the required actions might not have the right effect or impact or even might disturb the process.

The required time to perform all actions is the minimum operator response time. The maximum operator response time integrates the minimum operator response time with the allowable operator response time.

Update or review the alarm philosophy

The alarm philosophy is the place to specify how the alarm setpoints should be determined. It is also the place to indicate what the allowable response time for a given priority is estimated in your plant.

Inform everybody that alarms are there to be acknowledged before an operator action is performed.

An appropriate prioritization matrix for your plant should be chosen which is consistent with allowable and required operator response time and all stakeholders need to be aware of this matrix.

Also, the configuration of the reporting tool should be consistent with what is defined as timely in your alarm philosophy. For example: if the allowable operator response time is set to five minutes, any alarm that has a life cycle of less than five minutes could be considered as 'fleeting'. Similarly, if the maximum operator response time is set to one hour, any alarm that exists longer than one hour could be considered as 'long standing'.

Use a tool to measure the time the operator needs to act

Prior to updating the alarm setpoints of a given alarm system, the monitoring and assessment tool, eventually combined with historical data, could be used to assess the time the operator responds, the process deadtime and the effectiveness of the operator action.

Alarms that are cleared before any operator action is performed, or even before the operator acknowledges the alarm, are usually, if not always, notifications which should not alarm and hence be rationalized.

Document the estimated required operator response time

When using an appropriate alarm documentation tool, such as ProcessVue Guardian, the estimated required operator response time from the alarm rationalization process should be documented and be made available for the operator. During the life cycle of the alarm, this value should be updated as the alarm system evolves, and more data becomes available.

Avoid multiple alarms at the same time on the same console

Good alarm management aims at this goal. When a second alarm is annunciated during the time an operator is dealing with another, the operator gets distracted, the operator needs to deal with more than one item, which is error prone. See also the white paper on Alarm Performance KPIs.

Use the audit process to assess the allowable and required operator response time against the methods to define the alarm setpoint or trigger

The operator interviews during the audit process shall provide feed-back on how the timing of alarms is perceived. Do the operators assess the alarms as coming too early or too late? The operators should be asked if the maximum operator response time is sufficient to deal with the workload and the sanitary breaks the operator needs to take.

References

Endsley, M.R.: Towards a Theory of Situation Awareness in Dynamic Systems. Human Factors Journal 37(1), 32 – 64

EEMUA Publication 191 First Edition, Second Edition, Third Edition

ANSI/ISA 18.2 – 2009 and 2016
Lieven Dubois, Jean-Marie Forêt, Philippe Mack & Leen Ryckaert: Advanced logic for alarm and event processing: Methods to reduce the cognitive load for control room operators, Paper presented at the IFAC HMS Conference in Valenciennes, 2010
ISA 18.2 TR4, Advanced Alarming
ISA 18.2 TR6, Alarms in Batch Operations
ISA 18.2 WG8, TR8 under construction
IEC 62682 – 2015
MAC Solutions - ProcessVue Analyser (www.processvue.com)

About the author

Lieven Dubois (°1957) studied first electronic engineering and then software engineering in Belgium. He got involved in alarm management in the 1990s, where he introduced real-time expert systems to assist operators in dealing with abnormal situations.

Lieven started contributing to the ISA 18.2 Technical Report 4 on Advanced Alarming in 2009 and is now voting member of ISA 18.2 Alarm Management Committee. He was elected co-chair of Working Group 8 preparing a technical report on Alerts, Prompts, Notifications and Events.

Lieven was also involved in the preparation of the International Electro-Technical Committee (IEC) 62682 standard and the ISA 18.2 2016 edition. He is a qualified ISA IC39 alarm management course instructor.

Lieven is co-author of several papers about the application of real-time Artificial Intelligence technology, in particular to the domain of Alarm Management. He is a multilingual presenter on Alarm Management and Situational Awareness at several seminars, workshops and conferences, among others at the triennial IFAC HMS conferences in Valenciennes 2010 and Las Vegas, 2013.

Lieven is also co-author of the paper ‘Alarm Enforcement, or not’ with Ian Brown (2017) and author of the paper ‘Sense and Nonsense of Alarm system performance KPIs’ (2018)