Effective Resource Utilization for Alarm Management

Iman Izadi, Sirish L. Shah and Tongwen Chen

Abstract—In industrial plants, operators constantly receive a large number of alarms that are mostly false or nuisance. A majority of these alarms are generated by a small number of process variables known as bad actors. These bad actors are either poorly controlled which result in a lot of fluctuations; or their alarms are poorly configured. There are a number of products on the market and many methods in the literature that can identify and give recommendations to resolve bad actors. Nonetheless, although rectifying bad actors will significantly drop the alarm count, it is not nearly enough. To bring the performance of an alarm system within an acceptable range (given by a number of standards), we need to go further than fixing bad actors. This second step, if not more, is as difficult and time consuming as rectifying bad actors. In this paper we discuss and present methods and ideas to improve the performance of alarm systems beyond the bad actors.

I. INTRODUCTION

Industries reportedly lose billions of dollars annually due to unwanted accidents, equipment damage and unplanned shutdowns [1]. These incidents are generally caused by a variety of reasons from simple equipment malfunctions to major design flaws to avoidable human errors. Moreover, safety has grown to be the most important concern over all sectors of industry.

To assure an acceptable level of safety and to maximize profit margin, industrial plants are constantly monitored. Different monitoring techniques are practiced throughout industry to indicate any abnormality in the behavior of a plant, or detect any fault in the components of the plant. Fault detections and diagnosis have been and still are active fields of research both in academia and industry [6], [10].

After a fault or any other abnormality is detected in the plant, an alarm is raised to notify the operator about the unexpected situation. The operator then needs to carefully review the alarm and understand the source of the problem in order to take the necessary corrective action. Modern digital technology has made it very easy to implement various fault detection techniques and configure as many alarms as wanted. This has led to an increased number of alarms constantly generated, which unfortunately distract operators from their main duty [14]. As a result, alarm systems which were designed to make it easier to detect and solve plant problems in the first place, have themselves become a source of problem for operators and engineers alike.

Many attempts have been made recently to rectify the alarm situation in industry. A few standards have been developed on alarm systems [4], [9]. Various studies have also been done to set some guidelines on design and configuration of alarm systems [5], [7], [17]. Many companies too have developed software products that help plant authorities manage their alarm systems. They also offer a number of services aimed at making alarm systems more efficient and reliable by bringing down the alarm count.

In this paper, we first review the current situation of how alarms are handled in industry. Then we show why the current technology is not able to fully rectify the problems in alarm systems. To help this problem we propose a number of ideas that with further research and development can lead to better, more reliable and more automated alarm systems. The ideas are presented based on combined information from three different resources: process data, alarm data and process knowledge.

II. THE 80-20 RULE IN ALARM SYSTEMS

Industry standards (ISA 18.2 and EEMUA) suggest that, on average an operator should not receive more than 6 alarms per hour (one alarm every 10 minutes) during the normal operation of the plant [4], [9]. This number is recommended based on the fair assumption that, an operator needs about 10 minutes to promptly respond to each alarm. On the other hand, various surveys from different sectors of industry show that on average operators receive more than 30 alarms per hour even during normal plant behavior [11].

The interesting fact is that, most of these alarms are generated by only a handful of alarm variables. In other words, a very small percentage (20% or less) of alarm variables are responsible for a very large percentage (80% or more) of alarms (the 80–20 rule). These alarm variables are commonly known as bad actors. Often times even one bad actor can substantially increase the alarm count. We have seen extreme cases where only one alarm variable has generated more than 50% of the total alarm load.

As a common and very effective practice in industry, bad actors are reviewed on a regular basis, as it is expected...
resolving these bad actors could improve performance significantly. These facts have been directly addressed in alarm standards, e.g., in ISA 18.2 it is stated [9]:

“Relatively few individual alarms (e.g., 10 to 20 alarms) often produce a large percentage of the total alarm system load (e.g., 20% to 80%). The most frequent alarms should be reviewed at regular intervals (e.g., daily, weekly, or monthly). Substantial performance improvement can be made by addressing the most frequent alarms”.

One of the performance metrics that has been emphasized in alarm standards is that, the top 10 most frequent alarms should not contribute more than 5% to the total alarm load. In summary, no bad actors are acceptable.

Fortunately, current technology in alarm management (software products and services) is capable of identifying and addressing bad actors. Some of the major players in alarm industry offer products and services that are specifically designed to identify and attack bad actors. Addressing bad actors almost immediately improves the performance of alarm systems. In a recent industrial alarm rationalization practice, the alarm count dropped from about 60 alarms per hour to about 12 alarms per hour by only fixing bad actors.

Now, an important and critical question is: Can we bring performance of an alarm system within the desirable metrics (described by alarm standards), by only addressing bad actors? The answer is no. Addressing bad actors is the first step for improving alarm performance, but to comply with standards we need to look beyond the bad actors. This second step is often more challenging and requires more resources than the first.

Current status of industry suggests that, available technology is not capable of fully addressing the issues with alarm systems. Even if an industrial plant is able to comply with standards, it is usually achieved through a time and resource consuming alarm rationalization practice which is heavily based on experts knowledge rather than technology [16]. We believe that by developing better theoretical results and more advanced technology, the problem of alarm systems can be addressed and solved more effectively and efficiently; and this is not possible unless all available resources (process data, alarm data and process knowledge) are fully utilized. In the rest of this paper, we discuss these resources and what benefits they can offer in the design and analysis of alarm systems.

III. PROCESS DATA

Introduction of computer technology and advanced networks as well as decline in cost of instrumentation has changed process industry forever. Almost all the components in a modern plant are connected to some network nowadays, sending and receiving data constantly to/from other components. The data, whether in the form of an instant measure of a variable, or a message generated by a component, is used for different purposes including: control, monitoring, maintenance, and management. In most cases, data is stored in a database for future use as well.

In a modern plant, every variable that can be measured is measured. In a normal size plant we might have tens of thousands of variables measured and communicated over the network at a relatively fast rate (sub-second to a few seconds). Storing all these measurements requires huge amounts of memory and also makes it very difficult to manage and use the data effectively. Therefore, measurements are often stored at a much slower rate (every few tens of seconds). Stored data can be a compressed version of original data using some compression technique, or just a down-sampled version. Nonetheless, process variable measurements that are stored in a database (or process data as it is more conveniently known), is an extensive, rich and accessible source of data. Figure 1 illustrates a sample of process data extracted from a typical database.

Process data provides a comprehensive history of the plant making it a valuable resource for modeling, controller design, condition based monitoring, performance analysis and maintenance scheduling. This valuable resource, however, has not been fully utilized in the design of alarm systems. To the best knowledge of the authors, there have been very few research studies on the use of process data for alarm design [2], [8], [12], and even fewer commercial products. In this section we discuss how process data can be used for the design of alarm limits and alarm processing techniques.

A. Process Data and Alarm Design

Traditionally an alarm is defined in terms of a single variable exceeding some “normal” or “safe” range of operation (Figure 2). This is the simplest and most common practice in alarm generation. As a result the majority of alarms received by an operator are based on single variables going outside their corresponding alarm limits.

Consider the single-variable process data shown in Figure 3. By observing the behavior of this variable, we notice that, for about the first half of data, the variable remained constant with little variation. In the second half (from about
sample 1000), the variable shows a lot of change both in its magnitude and variability. Let us assume that the first half of data corresponds to the expected normal behavior. The second half of the data then shows some deviation from the normal behavior, possibly due to an unwanted underlying cause. Also assume that our objective is to notify the operator about the behavioral change by raising a timely alarm. In the ideal situation, no alarm should be raised during the normal behavior of the variable (the first half). The alarm then should be triggered as soon as the variable moves away from its normal state and should not be cleared until the abnormal behavior is rectified and the variable is back to normal. Needless to say, identifying normal and abnormal behavior of a variable is a subjective and non-trivial task that often requires additional knowledge about the process and variable. It remains one of challenges in design and analysis of alarm systems.

Design of an alarm framework for this variable is a standard classification problem. Many methods are available in the literature to solve this classification problem (a thorough study is given in [3]). The optimal solution is obtained by minimizing two miss-classification rates: false-positive (also known as type I error or false alarm) rate, and false negative (also known as type II error or missed alarm) rate. False alarm rate is a measure of alarms that are raised falsely during the normal behavior of the variable. Similarly, missed alarm rate measures alarms that are missed during the abnormal behavior of the variable. A general solution of this classification problem consists of a function (filter) and a threshold (alarm limit). For implementation, the filter is applied to the original data to obtain a filtered variable. The filtered variable is then compared to the alarm limit and if it exceeds the limits an alarm is raised.

This solution, although theoretically sound, has a few limitations for practical applications. First of all the designed filter is generally nonlinear. Therefore, the simple and trivial relation between the original and filtered data is more or less lost. This makes it difficult for an operator to understand the alarm and relate it to the behavior of the original variable. Secondly, the designed filter depends heavily on statistical distribution of process data. Therefore, for different process variables one might get totally different filters which makes it unrealistic for implementation. Moreover, many of distributed control systems (DCS) have limitations on implementing general filters. A more feasible solution is to restrict the type of filter to the most common ones. Moving average (MA), exponentially weighted moving average (EWMA) and cumulative sum (CUSUM) are examples of well known and commonly used filters in process monitoring and statistical quality control [12], [15].

As an alternative to filters, other processing techniques might be used for alarm generation. Here we briefly review deadbands and delay timers as the two most common alarm generation techniques.

### B. Deadband

When a deadband is configured for a variable, the alarm is raised, as usual, when the variable exceeds the threshold. However, the alarm is not cleared when the variable returns within the threshold. Instead, a tighter limit is used for clearing the alarm. Deadbands do not affect the way alarms are raised, but they make sure the alarm remains on longer than usual until the variable is within a tighter acceptable range.

Deadbands are the most common practice in alarm generation and have proven to be effective in reducing alarm chattering. A chattering alarm is defined as “an alarm that repeatedly transitions between the alarm state and the normal state in a short period of time” [9]. Alarm standards ISA 18.2 and EEMUA provide some guidelines on how much deadband should be used for each variable based on the nature of the variable. For instance for a level variable 5% deadband is suggested [4], [9].

Alarm deadband recommendations in the standards provide an acceptable initial value but for better performance, deadband should be fine tuned. Historical process data has an important role here as it can be used as a basis for deadband design. In [8] a method was proposed to design deadband based on the time series analysis of the process variable and its noise model. Another method to design optimal deadband was developed in [12] based on the receiver operating characteristic (ROC) curve. Despite these proposed techniques, the problem of deadband design is still open and requires further and deeper study.

### C. Delay Timer

Another widely used technique in alarm generation is the delay timer. Generally, as soon as a variable exceeds the threshold and enters the abnormal region, an alarm is raised. Similarly, as soon as the variable returns to the normal state the alarm is cleared. If delay timer is configured on a
variable, when the variable exceeds the threshold the alarm is not raised immediately. It is instead delayed for a number of samples and only if the variable remains in the abnormal state an alarm is raised. Similarly, the alarm is cleared only when the variable remains in the normal state for a number of samples.

Delay timers have also proven to be very effective in reducing the number of false and missed alarms. Similar to deadbands, standards suggest some guidelines on how much delay should be configured for different variables (for instance 60 seconds for a level variable) [4], [9]. These initial recommendations should also be fine tuned to ensure better performance of the alarm system. Historical process data can again serve as a basis for delay timer design as proposed in [12]. The design of delay timer and its effect on alarm performance is not yet well established and needs further study.

IV. ALARM DATA

In modern control and monitoring systems, alarms are generated within the DCS. An alarm is essentially a text message that is sent by the DCS to an output device (operator’s monitor, printer, log file, etc). In addition to alarms, DCSs send a wide range of other messages indicating different events in the plant. These events include return to normal (when a variable goes back to the normal state and the alarm is cleared), acknowledgment (when the operator acknowledges receipt of an alarm), operator action (when the operator takes a corrective action to rectify the alarm state) and other system related events.

Messages generated by different DCSs and other message generating devices are usually collected by another software. This software (often third party) parses and interprets each individual message from different devices and stores them in one database (known as the Alarms and Events (A&E) database). The data stored in this database (i.e., alarm data) can be extracted and analyzed for auditing, performance evaluation and maintenance purposes. Figure 4 illustrates a sample of alarm data.

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Alarm Id</th>
<th>Message Type</th>
<th>Value</th>
<th>Priority</th>
<th>Tag</th>
<th>Plant</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/6/2006 16:59:44.000</td>
<td>PUBH</td>
<td>OperatorAction</td>
<td>20.003</td>
<td>HIGH</td>
<td>1Q_237IC</td>
<td>Plant1B</td>
<td>Upgrading</td>
</tr>
<tr>
<td>11/6/2006 20:17:55.000</td>
<td>PUBLO</td>
<td>ReturnToNormal</td>
<td>199.526</td>
<td>HIGH</td>
<td>1Q_237IC</td>
<td>Plant1B</td>
<td>Upgrading</td>
</tr>
<tr>
<td>11/7/2006 09:39:11.000</td>
<td>PUBLO</td>
<td>Alarm</td>
<td>125.000</td>
<td>HIGH</td>
<td>22122</td>
<td>Plant12</td>
<td>Extraction</td>
</tr>
<tr>
<td>11/7/2006 09:42:48.000</td>
<td>PUBLO</td>
<td>Acknowledge</td>
<td>HIGH</td>
<td>22122</td>
<td>Plant12</td>
<td>Extraction</td>
<td></td>
</tr>
<tr>
<td>11/7/2006 10:08:11.000</td>
<td>PUBH</td>
<td>ReturnToNormal</td>
<td>178.466</td>
<td>HIGH</td>
<td>22122</td>
<td>Plant12</td>
<td>Extraction</td>
</tr>
<tr>
<td>11/7/2006 08:12:04.000</td>
<td>PUBH</td>
<td>OperatorAction</td>
<td>00</td>
<td>HIGHHIGH</td>
<td>22122</td>
<td>Plant12</td>
<td>Extraction</td>
</tr>
<tr>
<td>11/6/2006 16:28:06.000</td>
<td>PUBH</td>
<td>Alarm</td>
<td>00</td>
<td>HIGHHIGH</td>
<td>47_22M</td>
<td>Plant4</td>
<td>Power Plant</td>
</tr>
<tr>
<td>11/6/2006 16:28:22.000</td>
<td>PUBH</td>
<td>ReturnToNormal</td>
<td>00</td>
<td>HIGHHIGH</td>
<td>47_22M</td>
<td>Plant4</td>
<td>Power Plant</td>
</tr>
<tr>
<td>11/6/2006 16:29:31.000</td>
<td>PUBH</td>
<td>Alarm</td>
<td>00</td>
<td>HIGHHIGH</td>
<td>47_22M</td>
<td>Plant4</td>
<td>Power Plant</td>
</tr>
<tr>
<td>11/6/2006 16:35:31.000</td>
<td>PUBL</td>
<td>Acknowledge</td>
<td>236.02</td>
<td>HIGH</td>
<td>KIC_F17DB</td>
<td>Plant12</td>
<td>Pipeline</td>
</tr>
</tbody>
</table>

Fig. 4. A sample of alarm data obtained from database

Other information related to the alarm and the underlying cause is also included in the message. DCSs allow a number of different alarms to be configured for each variable. For instance, different messages are generated when the variable exceeds a high alarm limit, or it drops below the low alarm limit, or its rate of change is more than a threshold and so on. Each scenario constitutes a different alarm. Therefore, the number of possible alarms in a plant is far more than the number of process variables. Moreover there are some alarms in the system that are not generated by a process variable. These alarm, sometimes called system alarms, show the status of different components of the plant. For example, when a pump is jammed or does not start an alarm is generated to report the failure.

Alarm data is more complicated than process data. Unlike process data, alarm data and its format heavily depends on the hardware/software used to generate and collect messages.

The first challenge in dealing with alarm data is graphical representation. For process data, one can plot time trend of each individual variable. Alternatively, one might use different data compression methods (e.g., PCA) for graphical representation of a group of variables. This is possible due to the fact that, generally each and every process variable has a value at every time instant. For alarm data, on the other hand, the situation is different as for each alarm variable, the alarms are sparsely and non-uniformly distributed over time. Simple graphical representations, mostly in the form of bar or pie charts, are readily available in commercial Alarm Management software. Alarm count by tag, alarm tag by operator and alarm distribution over time are examples of these trivial charts. More advanced visualization techniques such as High Density Alarm Plot (HDAP) and Alarm Correlation Color Map (ACCM) have been recently developed [13]. These new methods have the advantage of simultaneously visualizing a group of alarm variables which make it possible to observe similarity and correlation between different alarm variables.

In addition to visualization, alarm data is a rich source of information for analysis and design of alarm systems. In fact, all performance metrics given in alarm standards ISA 18.2 and EEMUA (e.g., alarm count per hour, alarm chattering, and peak alarm rate) are calculated based on alarm data.
These performance metrics provide a simple and meaningful indicator of the status of the alarm system. But they are not suitable for efficient and quantitative alarm design (e.g., they are not suitable to be used as the objective function in an optimization problem). More comprehensive performance indices are required to optimally design an alarm system.

Data mining of alarm data is another interesting topic. During a plant upset, often the alarm count increases drastically to hundreds or even thousands of alarms per hour. To identify the root cause of the upset, the operator is required to go through all the alarms and separate the more significant ones. Moreover, after each incident, a team of experts review the log files to diagnose the source of the problem for future prevention. These types of analysis are now heavily based on operator’s experience and expert’s knowledge. Data mining and pattern recognition techniques can provide a tool to extract meaningful data during an alarm flood. This is useful both for online implementation for operators and off-line analysis.

Another challenge is how to use alarm data for the design of alarm systems. If performance measures are not within the acceptable range, how can the alarm system be modified to achieve the standards? For example, if an alarm is chattering what technique could be used to reduce the chattering? Should the alarm generation framework be modified (e.g., by changing deadband, delay timer, or alarm limit)? Or should the alarm configuration be modified (for instance, by changing alarm priorities or removing the alarm)? Strategies such as delay timers could also be configured on an alarm, i.e., if the alarm is raised it should not be cleared for a particular period of time even if the underlying process variable returns to normal.

V. PROCESS KNOWLEDGE

Process knowledge is the term we reserve for all the information and knowledge about a plant that is based on the laws of nature and is not time dependent. Process knowledge can be divided into three categories:

- **Engineering knowledge** captures the relationships between different parts of a process based on laws of physics. Mass and energy balance, laws of thermodynamics, Newton’s laws of motion and Kirchhoff’s circuit laws are example of laws that help us build our engineering knowledge about the plant. The engineering knowledge is often represented by mathematical models of the process.

- **Operator experience**. Operators develop a very good understanding of the plant behavior over time. They can explain actions and reactions between different parts of the plant based on sheer experience, even when the engineers fail to do so.

- **Connectivity data** is the information of physical connections between different components of a process, which captures the flow of material and information throughout the plant. Connectivity data is often represented graphically by process flow sheets and process and instrumentation diagrams (P&ID). Figure 5 shows a sample of connectivity data.

Process knowledge is the main source of information for initial configuration of an alarm system as well as the alarm rationalization process. In an alarm rationalization process, a team of experts including process engineers, experienced operators and a moderator with expertise in alarm systems, review all the alarms in the plant. In addition to their knowledge and experience, they rely heavily on connectivity data to reconfigure the entire alarm system. The alarm rationalization is a costly and time/resource consuming process, mostly because it is completely manual. Any solution that automates the alarm rationalization process or at least a part of it would be considered a breakthrough in industry.

Process knowledge, specifically connectivity data, can be used to investigate and explain how an abnormality is propagated through a process. This information can then be used to predict how an abnormality can trigger a series of alarms. We can then use this to group alarms during a plant upset and track back a series of alarms to the original cause of abnormality (root cause analysis). Moreover, important loops and variables in a process can be identified from connectivity data for better alarm configuration.

Process knowledge is a far more complicated resource than process and alarm data. Even the mere capture of process knowledge by a computer program is very challenging, let alone the automatic use of this resource for alarm analysis and design.

We can translate connectivity data and process knowledge to more qualitative graphs such as a signed directed graph (SDG). SDGs qualitatively capture the cause and effect relationships (causality effects) between different process variables. They show how increasing one variable can increase or decrease other variables. SDGs have proven to be useful in alarm analysis and can help with configuration of alarm systems [18]. Nonetheless, the procedure of obtaining an SDG from connectivity data is still manual and remains a challenge.

VI. COMBINING RESOURCES

Combining the three resources mentioned in the previous sections provide even more information about the alarm system. Here we propose a few ideas:
A. Combining Process and Alarm Data

Alarm data can be used to identify normal and abnormal behavior of a process variable, which is one of the challenges in using process data for alarm design. Based on alarm data, we can mark the part of process data that generates an excessively large number of alarms as abnormal data. Another interesting topic is how correlation between process data can be related to similarity between alarm data. In other words, based on correlation structure of process data and similarity of alarm data, can redundant alarms be identified?

From another perspective, process data can be used to evaluate the performance of an alarm system. Measures such as false alarm rate, missed alarm rate and detection delay, which are based on process data, can be combined with alarm performance measures to construct more comprehensive evaluation criteria. Combined visualization of process data and alarm data has also proven to be extremely useful in verifying an alarm system and is also another interesting research topic.

B. Combining Process Data and Process Knowledge

Process knowledge is the most reliable way to identify normal and abnormal behavior of a variable. Normality is a subjective term and process knowledge can put process data in a meaningful perspective making it easier to mark normal and abnormal behavior.

As mentioned before, building a qualitative graphical representation of a process (e.g., SDG) is manual and time consuming. However, investigating the correlation structure of process data can highlight the causality relations between process variables to some extent, which in turn makes it easier and more automated to construct a connectivity graph [18]. The correlation structure of process data is also helpful in verifying, modifying and validating a qualitative graph such as SDG.

C. Combining Alarm Data and Process Knowledge

Similar to process data, alarm data may also be useful in constructing and verifying a qualitative connectivity graph and reducing the manual burden. This is possible by studying the similarities between alarm variables and identifying causality relations in alarm data.

VII. CONCLUDING REMARKS

In this paper we have discussed the current status of alarm systems in industry. The situation now, although much better than a few years ago, is not yet satisfactory. To reach the acceptable point and comply with alarm standards, new theoretical results and more advanced design and analysis technologies are required. Effective and complete utilization of different resources (process data, alarm data and process knowledge) is a key in achieving this goal.

Process data is the simplest and most available source of information for a plant. However, this rich resource has not yet been utilized in alarm design. Process data can be used to optimally design alarm processing technique (filters, deadbands and delay timers) as well as alarm limits.

Alarm data is a more complicated source of information that is currently used by alarm standards and commercial Alarm Management software to evaluate the performance of an alarm system. For alarm data, visualization, data mining, alarm analysis and alarm design are some of the topics that have not been extensively studied.

Process knowledge is the most complicated resource in terms of how to automatically capture the information. It would be used extensively in initial design of alarm systems and alarm rationalization. However, most of these procedures are still manual and require immense resources to be implemented properly. Systematic use of process knowledge and connectivity data in alarm analysis and design is still a challenging topic. Any progress in this field brings us one step closer to automating the alarm rationalization process.

Many of the challenges and research topics mentioned in this paper are now under investigation. Nevertheless, from the theoretical point of view, the area of analysis and design of alarm systems is still widely open and untouched.

REFERENCES