12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

Identification of Most Critical Alarms for Alarm Flood Reduction

Md Habibur Rahaman, Haniyeh Seyed Alinezhad, Tongwen Chen

Department of Electrical & Computer Engineering, University of Alberta, Edmonton, Alberta T6G 1H9, Canada
Email: mdhabibu@ualberta.ca, h.aalinezhad92@gmail.com, tchen@ualberta.ca

Abstract: In complex processes, the activation of a single alarm can trigger a cascade of consequences that affect multiple interconnected components. As a result, the number of active alarms can increase rapidly. This sudden surge in alarms is often referred to as an alarm flood. Alarm floods are a common source of operational burden for operators, overwhelming them with a high volume of alarm notifications. If critical alarms are not promptly and accurately identified, decision-making processes can be undermined. This paper addresses these challenges by introducing a novel approach for identifying and prioritizing critical alarms from each alarm flood. Hidden Markov models are employed to construct a likelihood matrix that reveals the relationships among alarm variables, and identifies the most critical alarm from a directed acyclic graph. Case studies are conducted using a vinyl acetate monomer simulator to demonstrate the effectiveness of the proposed approach. The results highlight accurate identification and prioritization of critical alarms, enabling operators to focus on the most important process abnormalities.

Keywords: Process data, alarm data, alarm floods, alarm ranking, critical alarms, and fault diagnosis.

1. INTRODUCTION

Modern industrial systems are often characterized by their large scale, involving numerous components and variables, which create a complex system topology. This complexity can result in fault propagation, which can cause a problem called alarm flood. Identifying critical alarms for each alarm flood is essential for effective operator response. As industrial facilities expand, detecting the causes of alarm propagation becomes more challenging. Managing alarm systems is crucial for ensuring safety since alarms serve as indicators of process abnormalities (Alinezhad et al., 2022). However, the growing complexity of these systems has led to more alarms during alarm floods, making it difficult for operators to prioritize responses. Relying solely on human operators becomes increasingly difficult as industrial systems evolve. This necessitates the integration of automated tools with advanced analytical and visualization capabilities to interpret complex data patterns (Alinezhad et al., 2023).

The development of effective alarm systems has been a topic of significant research interest in industrial process control. Over the years, various approaches and techniques have been proposed to address the challenges associated with the increasing number of alarms and the need for operators to efficiently handle them. One common aspect of alarm system research is the identification of most critical alarms, which are alarms that require immediate attention and action to prevent adverse consequences. Several studies have focused on developing methods to prioritize alarms based on their impact on process safety and operational performance, which are reviewed as follows.

One prominent approach is the use of data-driven techniques to analyze alarm data and extract meaningful insights. For instance, Abele et al. (2013) proposed a method combining statistical analysis and machine learning to identify most critical alarms in a chemical plant. They used a clustering method to group similar alarms and applied a decision tree algorithm for prioritization. In addition to data-driven approaches, researchers have explored the use of knowledge-based methods for alarm identification. Fukui et al., (1989) proposed a knowledge-based approach using process knowledge and expert rules for alarm prioritization, improving most critical alarms identification and reducing false alarms. Another vital aspect is visualizing alarm patterns and dependencies in alarm systems. Jiang et al., (2023) introduced a graphical method using directed acyclic graph (DAG) plots to visualize alarm hierarchy and propagation paths, aiding operators in understanding the alarm system and identifying most critical alarms based on their positions within the graph. Rodrigo et al. (2016) presented a comprehensive approach for identifying the root causes of alarm floods in large-scale industrial systems. Noroozifar et al., (2019) demonstrated the application of alarm data for root cause analysis of process faults, providing valuable insights into the causes of process faults and data trends, and ultimately enhancing the reliability of industrial systems. Zhang et al., (2022) proposed a novel causal fusion method was combined with process topologies and alarm data to identify alarm root causes. However, a potential drawback is the complexity of the method and computational demands. Recently, hidden Markov models are utilized to diagnose plant alarms early by analyzing fault propagation paths during abnormal conditions (Venkidasalapathy et al., 2019 and Hu et al., 2018). However, these studies have two potential limitations: the model accuracy and the suitability for handling multi-fault situations.

Although progress has been made in fault diagnosis, complex industrial processes require more comprehensive and effective approaches. A key challenge is classifying faults, diagnosing them, and identifying most critical alarms to provide operator assistance in case of alarm floods. Unlike conventional methods that assign priority based on the impact of specific process variables, this study employs a data-driven analysis to identify critical alarms using a simple method and with low computation time.

This paper is structured as follows: Section 2 gives the problem statement of this research. Section 3 presents the methodology and the various steps employed to achieve the objectives. In Section 4, a case study is presented to validate the effectiveness of the proposed methodology. Finally, Section 5 provides concluding remarks.

2. PROBLEM STATEMENT

An effective alarm system is essential to ensure the secure and smooth operation of complex industrial processes. This research aims to enhance alarm system management and decision support for operators by developing a comprehensive approach based on alarm prioritization. Let a set of unique alarm variables in a process system be represented by $\mathcal{A} = \{a_n : n = 1,2,\dots,|\mathcal{A}|\}$, where $|\mathcal{A}|$ is the total number of alarm variables. An individual alarm variable is assigned with a criticality level, denoted by $\mathbb{C}(a_n)$, which quantifies its importance or impact on process safety and performance. The criticality levels are classified into three categories: Low, High, and Critical. An alarm flood can be identified using the concept of alarm rates with predefined thresholds within a time bin of size $T = 10$ min (ANSI/ISA-18.2, 2016). Each alarm flood can be denoted as ϕ_k, where k is the index of the alarm flood, with starting and ending time denoted by t_s^k and t_e^k, respectively. The set of alarms in ϕ_k is denoted as $\mathcal{F}_k$=$\{a_x \in \mathcal{A}, x = 1,2,\dots,|\mathcal{F}_k|\}$, where $|\mathcal{F}_k|$ represents the total number of unique alarm variables in ϕ_k, such that alarm a_x occurred at least once in the time period $[t_s^k - T, t_e^k]$. It is obvious that $\mathcal{F}_k \subseteq \mathcal{A}$.

We aim to construct a likelihood matrix, denoted by $\mathcal{L}_{\mathcal{M}}$, to represent the probability of transitioning from one alarm variable to another in an alarm flood. Then, a directed acyclic graph plot can be created for each alarm flood using the constructed $\mathcal{L}_{\mathcal{M}}$. As a result, the most critical alarm, denoted by a_k^*, can be determined by analyzing the directed acyclic graph plot and counting the number of outgoing edges from each alarm a_x, denoted as $\xi(a_x)$. The alarm with the maximum number of outgoing edges can be considered the most critical alarm for that alarm flood which can be represented as follows:

$$a_k^* = \underset{a_x}{\operatorname{argmax}}\{\xi(a_x)\} \tag{1}$$

This analysis contributes to enhancing operator decision support by enabling the prioritization of critical alarms, a_k^* for alarm flood ϕ_k and thus improving the overall management of alarm systems in complex industrial environments.

3. METHODOLOGY

This section presents the proposed methodology, which encompasses a series of interconnected steps aimed at addressing the complexities inherent in industrial processes. The detailed architecture of the methodology is shown in Fig. 1.

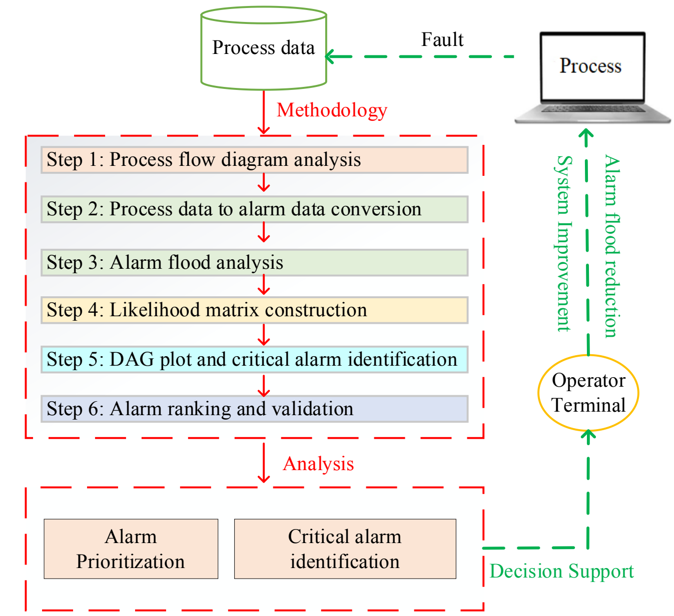

Figure 1. The detailed architecture of the proposed methodology.

This is intended to improve the decision-making capabilities of operators and reduce alarm floods. In the following, a detailed explanation of the proposed methodology is provided.

3.1. Process flow diagram analysis

The graphical representation of a complex process topology is shown in Fig. 2. The analysis of the process flow diagram involves examining the diagram to identify the process units and several distinct process elements (nodes) connected to each unit. The G_U represents the process topology, where $U = \{u_1, u_2, \dots, u_g, \dots |U|\}$ indicates unit (or plant area) set, u_g is the gth unit, $|U|$ is the total number of units; E_v is the undirected edge set. Here, $G_U = (U, E_v)$ is represented by the process topological adjacency matrix $\mathcal{M}$, i.e.,

$$\mathcal{M}_{gh} = \begin{cases} 1, & \text{if } (u_g, u_h) \in E_v \\ 0, & \text{otherwise} \end{cases}, \tag{2}$$

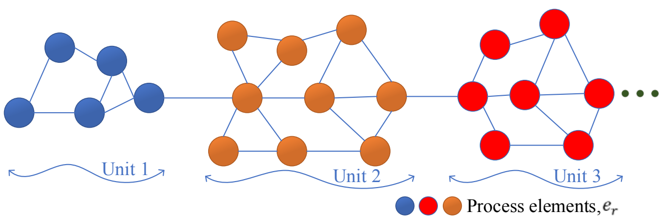

Figure 2. An example of a complex process topology.

where, u_g, and u_h represent two individual edge connecting units. $\mathcal{M}_{gh}$=1 represents that the gth row u_g connects with the

hth column u_h. The matrix $\mathcal{M}$ is always symmetric, i.e., $\mathcal{M}_{gh} = \mathcal{M}_{hg}$. The relationship between process units (u_g) and process elements, denoted as e_r, is hierarchical (Zhang et al., 2022) and depends on various factors such as material flow, energy transfer, and control logic. In real industrial systems, there are typically a few units corresponding to different processes, and each unit comprises multiple process elements. Consider $E = \{e_r : r = 1,2,\dots,|E|\}$ representing the set of distinct process elements, where $|E|$ is the total number of process elements. A hierarchical mapping matrix, $\mathcal{M}_{UE}$, of dimensions $|U| \times |E|$, can be created to represent the relationship between units and process elements. Each entry of $\mathcal{M}_{UE}$ will indicate whether unit u_g contains process elements e_r. Similarly, an alarm mapping matrix, $\mathcal{M}_{\mathcal{A}}$, of dimensions $|E| \times |\mathcal{A}|$, can be created to represent the relationship between process elements and alarms. Each entry in $\mathcal{M}_{\mathcal{A}}$ will indicate whether process element e_r generates alarm a_n or not. Each entry in these matrices signifies the presence or absence of a connection, with a value of "1" indicating a connection and "0" indicating no direct relationship.

Following the above procedure, the process elements and units associated with the analyzed alarms are identified. This information can be used to validate the identified most critical alarm responsible for specific faults within interconnected systems.

3.2. Process data to alarm data conversion and formation of alarm floods

One of the most challenging problems in industrial alarm management revolves around the occurrence of alarm floods. These floods often result from the propagation of process abnormalities (Rodrigo et al., 2016). The ISA-18.2 standard (ANSI/ISA-18.2, 2016) provides a concise definition of an alarm flood as "a condition in which the alarm rate exceeds the operator's ability to effectively handle." It further recommends a benchmark threshold of 10 alarms per 10 minutes per operator to identify the presence of alarm floods. The detailed conversion procedure of process data to alarm data is mentioned in (Parvez et al., 2022). Consider the signal form of an alarm $a_n \in \mathcal{A}$ given by

$$\tilde{x}_{a_n}(t) = \begin{cases} 1 & \text{if } p_v(t) \notin P_v \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

where P_v indicates the normal operating specification of a process variable $p_v(t)$. The signal form of alarm occurrence of $a_n \in \mathcal{A}$ is

$$x'_{a_n}(t) = \begin{cases} 1 & \text{if } \tilde{x}_{a_n}(t-1) = 0 \text{ and } \tilde{x}_{a_n}(t) = 1 \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

Identifying the existence of an alarm flood can be based on the alarm rate after removing the chattering alarms (Cheng et al., 2013), which is a widely used key index to evaluate the performance of an alarm system. It represents the alarm count over a time period of certain length. Mathematically, the alarm rate $\alpha(t)$ at time instant t in the past time bin of size T is defined as

$$\alpha(t) = \sum_{n=1}^{|\mathcal{A}|} \sum_{g=t-T+1}^{t} x'_{a_n}(g). \quad (5)$$

The start and end time of an alarm flood can be identified by comparing $\alpha(t)$ with predefined threshold, δ. An indexing variable ϕ can be denoted as the presence of alarm floods, which is defined as

$$\phi(t) = \begin{cases} 1 & \text{if } \alpha(t) \geq \delta_s, \text{and } \phi(t-1) = 0 \\ 0 & \text{if } \alpha(t) < \delta_e, \text{and } \phi(t-1) = 1, \\ \phi(t-1) & \text{otherwise} \end{cases} \quad (6)$$

where "1" and "0" represent the presence and absence of alarm floods, respectively. The initial sample $\phi(0)$ is set to be 0. Based on the ISA-18.2 standards (ANSI/ISA-18.2, 2016), the benchmark thresholds to identify the start and end of alarm flood are ten and five alarms over a 10-min period for each operator. Thus, two thresholds in (6) are δ_s=10 and $\delta_e = 5$ based on a time bin of size T = 10 min in (5). Finally, in a multiple alarm flood scenario, each alarm flood can be identified as ϕ_k, where k is the index of alarm floods.

3.3. Alarm flood analysis

In this subsection, we introduce two fundamental analyses of alarm floods: the determination of individual alarm flood durations and the process of forming alarm flood clusters. The variable ϕ_k is used to find the start and end time stamps of each alarm flood: An alarm flood is said to begin at time instant t_s^k if

$$\phi_k(t_s^k) = 1 \text{ and } \phi(t_s^k - 1) = 0 \quad (7)$$

which indicates the alarm rate $\alpha(t)$ reaching the threshold δ_s over a 10-min period, and end at time instant t_e^k if

$$\phi_k(t_e^k) = 0 \text{ and } \phi(t_e^k - 1) = 1 \quad (8)$$

which indicates the alarm rate $\alpha(t)$ dropping below the threshold δ_e over a 10-min period.

Once the alarm flood time intervals are isolated and the alarm flood sequences are built, similar alarm flood sequences are clustered using sequence pattern matching. A similarity matrix is built using pairwise similarity indices obtained from the modified Smith-Waterman algorithm (Cheng et al., 2013). Finally, agglomerative hierarchical clustering is applied to cluster the alarm flood sequences. The detailed steps were mentioned by Rodrigo et al., (2016). Alarm flood clusters can be denoted as $\mathcal{C} = \{c_d : d = 1,2,\dots,|\mathcal{C}|\}$, where $|\mathcal{C}|$ is the total number of clusters. If multiple alarms are associated with one alarm flood cluster, there may be one or two alarms that trigger other alarms, resulting in the propagation of abnormalities within interconnected process systems. Consequently, alarm flood cluster analysis can serve as a means to validate the identification of the most critical alarm.

3.4. Likelihood matrix construction

A likelihood matrix, denoted as $\mathcal{L}_{\mathcal{M}}$ can be obtained using a Markov chain model (Venkidasalapathy et al., 2022) that focused on the sequential relationships of alarms within individual alarm flood. In the context of Markov chain model, each unique alarm variable within an alarm flood represents a state in the Markov chain. Hereby, $|\mathcal{F}_k|$ is the total number of unique alarm variables in the kth alarm flood, that represents the number of Markov chain states. The number of times each alarm variable transitions to another alarm variable within the same alarm flood can be calculated as $\rho(a_i \to a_j)$ where a_i and a_j are alarm variables. $\mathcal{L}_\mathcal{M}$ serves as a representation of transition probabilities of alarm variable pairs. For each pair of alarm variables a_i and a_j, calculation of the transition probability is conducted as follows:

$$\mathcal{P}(a_i \to a_j) = \frac{\rho(a_i \to a_j)}{|\mathcal{F}_k|}. \quad (9)$$

The calculated transition probabilities can be used to construct the element of a matrix, $\mathcal{L}_\mathcal{M}$ as

$$\mathcal{L}_\mathcal{M}(i,j) = \begin{cases} \mathcal{P}(a_i \to a_j) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}. \quad (10)$$

It quantifies the sequential dependencies between alarms within each alarm flood, representing the fundamental concept of sequential dependency analysis. However, several factors, such as inaccurate records and missed alarms in dynamic processes, can significantly reduce the model accuracy.

3.5. Application of hidden Markov models

A hidden Markov model (HMM) is a probabilistic model that stands as a well-established tool for recognizing patterns and analyzing sequences in various fields. The fundamental theoretical explanation about the model parameters is discussed in Venkidasalapathy et al., 2022. In this subsection, we delve into the core components of the HMM and how they contribute to uncovering meaningful insights within complex data sequences. In an HMM, a sequence of unique alarm variables belonging to each alarm flood is taken into account. The set of observed alarm variable in each alarm flood can be denoted as $\mathcal{O} = \{o_T : T = 1,2,\dots,|\mathcal{O}|\}$, where $|\mathcal{O}|$ is the total number of observed alarm variables. Similarly, the set of hidden states can be denoted as $\mathcal{Q} = \{q_T : T = 1,2,\dots,|\mathcal{Q}|\}$, where $|\mathcal{Q}|$ is the total number of hidden states. The HMM comprises two essential sets of parameters, referred to as "transition probabilities" and "emission probabilities."

Transition probability: The transition probability, denoted as $\mathcal{T}_{ij}$, is the likelihood of transitioning from alarm a_i to alarm a_j within each alarm flood. The estimation of the element of $\mathcal{T}_{ij}$ can be obtained as follows:

$$\mathcal{T}_{ij} = \begin{cases} \mathcal{P}(q_{t+1} = a_j | q_t = a_i) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}. \quad (11)$$

Emission probability: The emission probability, denoted as $\mathcal{E}_{ij}$, is the probability of observing alarm o_j given state a_j at time step t within each alarm flood. The estimation of the element $\mathcal{E}_{ij}$ can be obtained as follows:

$$\mathcal{E}_{ij} = \begin{cases} \mathcal{P}(o_t = o_j | q_t = a_i) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}. \quad (12)$$

The details of the parameter estimation procedure can be found in Venkidasalapathy et al., 2022. Finally, a modified likelihood matrix is introduced based on (11) and (12), where each matrix element is defined as follows:

$$\mathcal{L}_\mathcal{M}(i,j) = \begin{cases} \mathcal{T}_{ij} \times \mathcal{E}_{ij} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}. \quad (13)$$

This matrix $\mathcal{L}_\mathcal{M}$ combines the joint influence of transition probabilities and emission probabilities for each pair of alarm variables. The $\mathcal{L}_\mathcal{M}$ is utilized to construct a DAG plot for each alarm flood, providing a visual representation of the influence relationships among the alarm variables.

3.6. DAG plot and critical alarm identification

The likelihood matrix is utilized to construct a DAG plot for each alarm flood, providing a visual representation of the influence relationships among the alarm variables. The DAG represents the alarm variables as nodes, and the directed edges indicate the influence between variables. Through visual examination of the constructed DAG, significant alarms can be identified, and the propagation of the influence of alarm variables throughout the network of alarm variables can be observed. Consider G as the DAG representing the influence relationships within one alarm flood. Each alarm variable is represented as a node in G. An edge from node i to node j in G indicates that a_i influences a_j, where the direction and strength of influence are captured from the elements of $\mathcal{L}_\mathcal{M}$.

3.7. Alarm ranking and validation

The total standing time (*TST*) can be calculated from alarm and event (A&E) log database for each unique alarm variable. Consider t_a^q and t_r^q as the timestamps of the qth occurrence and return to normal states of a unique alarm. The *TST* of the qth alarm occurrence for a specific alarm variable is denoted as $\mathcal{D}^q$ which can be calculated as

$$\mathcal{D}^q = t_r^q - t_a^q. \quad (14)$$

If $\mathcal{R}$ represents the total number of occurrences of the specific alarm variable, then the *TST* can be calculated by taking the summation of all occurrences for that alarm variable as follows:

$$TST = \sum_{q=1}^{\mathcal{R}} \mathcal{D}^q. \quad (15)$$

According to Noroozifar et al., 2019, if multiple alarms are annunciated simultaneously, then an alarm with a longer *TST* is considered a critical alarm. Therefore, all alarms within the A&E log can be ranked based on their *TST,* which can be used to validate the identified most critical alarm.

4. CASE STUDY

A case study is presented to validate the proposed method. A well-known process simulator named vinyl acetate monomer (VAM) simulator is used in this case study. Details of the operation and introduction of various units were discussed in Machida et al., 2016. The basic process units are the reactor, column, decanter, separator, compressor, absorber, buffer tank, and raw materials feed. An adjacency matrix is constructed from the process flow diagram based on (2). In the simulation steps, steady state data is extracted from the standard steady state simulation for a certain time. Several malfunctions are introduced in the VAM process and the responses are observed. Each disturbance or fault is set to occur only once, but it repeats frequently over the given duration. Finally, the faulty process data is extracted for further analysis. Process data usually consist of timestamps and measurement values for all process variables. The extracted process data is converted to alarm data by following the procedure mentioned in Yang et al., 2020, where a total of 124 unique alarm tags are identified. An example of A&E log data is presented in Table 1.

Table 1: A&E log data

Timestamp	Tag	Unit	Identifier	Priority	Status
03:08:36 AM	PDI501	Column	PVLO	Low	ALM
03:08:39 AM	PDI501	Column	PVLO	Low	RTN
03:08:51 AM	PDI501	Column	PVLO	Low	ALM
03:10:01 AM	PDI501	Column	PVLO	Low	RTN

As shown in Table 1, the timestamp indicates the time of occurrence. The tag name comprises an asset name and an identifier, for example, PDI501.PVLO. The identifier indicates the level of abnormality compared to a normal operation threshold. The priority column indicates the criticality level, $\mathbb{C}(a_n)$, which quantifies its importance or impact on process safety and performance such as Low, High, or Critical. Finally, the status represents the state of alarms, such as alarm state (ALM) or return to normal state (RTN). The formation and analysis of alarm floods are carried out following the procedures outlined in Sections 3.2 and 3.3 respectively. Chattering alarms are eliminated using a 30-second off-delay timer. As shown in Table 2, a total of 23 alarm floods are identified from the A&E log, where each alarm flood (AF) is considered as a distinct fault.

Table 2: Alarm flood analysis

AF ID	Duration	# of ALM's	# of Tags	Fault Label
1	0h-13m-58s	22	12	1
2	0h-15m-6s	21	16	2
3	0h-17m-5s	21	15	3
.	.	.	.	.
.	.	.	.	.
22	0h-16m-35s	18	10	22
23	0h-12m-27s	22	09	23

For demonstration purposes, the most critical alarm will be identified from the first alarm flood. The details of the first alarm flood, extracted from the A&E log, are provided in Table 3. The data is visually represented in Fig. 3.

Table 3: Alarm flood 1

Alarm Tag & ID	Time	Priority
QI404.PVHI	03:12:11	High
TP201PV(6).PVHI	03:13:12	Low
QI404.PVHI	03:13:35	High
.	.	.
.	.	.
QI404.PVHI	03:24:54	High
TP502PV(10).PVHI	03:26:06	Low

A total number of 22 alarm variables are found in the first alar flood, among which a total of 12 alarm variables are unique.

The MATLAB tool "HMMestimate" is used to estimate the HMM parameters presented in Section 3.5 (Venkidasalapathy et al., 2022).

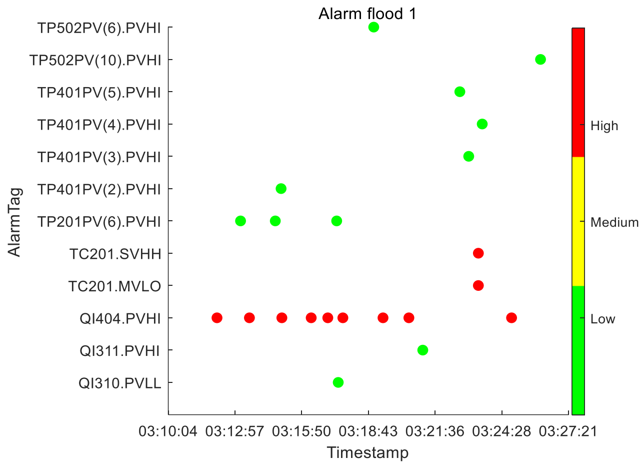

Figure 3. First alarm flood.

Once the transition and emission probabilities are obtained, a likelihood matrix is constructed. Since the number of unique alarm tags are 12 for the first alarm flood, the size of the likelihood matrices is 12×12. These matrices establish dependency relationships among alarm tags, which are treated as vertices of a DAG, and their associated weight represents the strength of dependency. The DAG plots for the first alarm flood is presented in Fig. 4. From Fig. 4, it is evident that the alarm tag "QI404.PVHI" and "TP201PV(6).PVHI" hold significant influence and represent potential critical alarms for the first alarm flood.

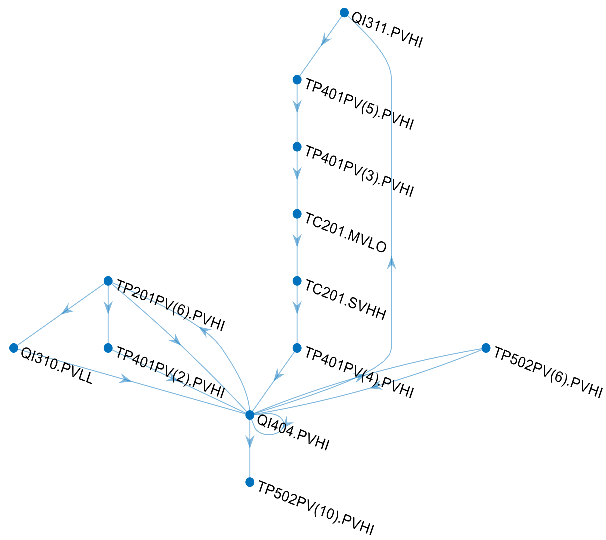

Figure 4. DAG plot for the first alarm flood

Based on Noroozifar et al., 2019, when multiple alarms occur simultaneously, the potential criticality of an alarm tag can be determined by comparing their TST. For instance, a longer TST suggests a potentially more critical alarm tag. The summary of the TST for all alarm tags from the main dataset is depicted in Fig. 5. This summary provides evidence that the TST of "QI404.PVHI" is notably higher than that of

"TP201PV(6).PVHI". This observation ensures the high criticality of "QI404.PVHI".

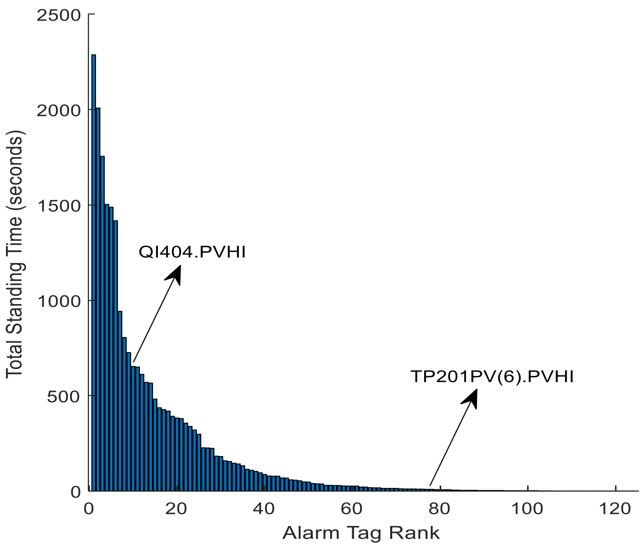

Figure 5. Alarm tags rank based on *TST*

Constructing the adjacency matrix from the process flow diagram involves defining a set of units, denoted as U = $\{\mathcal{u}_1, \mathcal{u}_2, \ldots, \mathcal{u}_7\}$, where each unit is represented by labels like "Absorber", "Buffer Tank", "Column", "Decanter", "Raw material feed", "Reactor", and "Separator and compressor". If two units share common alarm tags that satisfy the conditions outlined in Section 3.1, then an edge exists between them. This relationship can be mathematically represented as $\mathcal{M}_{gh}$= 1; otherwise, $\mathcal{M}_{gh}$ is set to 0. Simultaneously, the hierarchical mapping matrix, $\mathcal{M}_{UE}$, and the alarm mapping matrix, $\mathcal{M}_{\mathcal{A}}$, are also examined. Finally, the total number of alarms per process unit is determined. Such alarms can trigger other alarms, leading to the propagation of abnormalities within interconnected process systems. The identified critical alarm can be further investigated manually using the number of alarms within the interconnected process units, confirming the most critical alarm, a_k^* for a specific alarm flood.

5. CONCLUSION

In this paper, a data-driven approach was presented for identifying the most critical alarm for each alarm flood within complex process systems. By utilizing available alarm data and employing various analyses, the proposed method successfully pinpointed the alarms with the highest impact on overall system behavior. The proposed methodology offered a systematic and efficient approach to prioritize alarms, focusing on those that require immediate attention. Through the application of HMMs, the method identifies alarms displaying significant dependencies and influences within the system. The obtained results offered valuable insights into critical alarms and their relationships, empowering operators to make informed decisions and proactively address potential faults or abnormalities. Although this study achieved considerable progress in critical alarm identification, there exist several avenues for further research and improvement. One potential future direction involves integrating machine learning techniques and anomaly detection algorithms to enhance the accuracy of critical alarm identification. By incorporating real-time data and continuously updating the analysis, critical alarm identification can become more adaptive and responsive to evolving system conditions.

REFERENCES

Alinezhad, H. S., Roohi, M.H., and Chen, T. (2022). A Review of Alarm Root Cause Analysis in Process Industries: Common Methods, Recent Research Status and Challenges. *Chemical Engineering Research and Design,* 188, 846-860.

Alinezhad, H. S., Shang, J. and Chen, T. (2023). Open Set Online Classification of Industrial Alarm Floods with Alarm Ranking. *IEEE Transactions on Instrumentation and Measurement,* 72, 1–11.

Abele, L., Anic, M., Gutmann, T., Folmer, J., Kleinsteuber, M., and Vogel-Heuser, B. (2013). Combining Knowledge Modeling and Machine Learning for Alarm Root Cause Analysis. *IFAC Proceedings Volume,s* 46(9) 1843–1848.

Fukui, Y., and Masuzawa, M. (1989). Knowledge-based Approach to Intelligent Alarms. *Journal of Clinical Monitoring,* 5(3), 211–216.

Jiang, W., and Bai, Y. (2023). APGNN: Alarm Propagation Graph Neural Network for Fault Detection and Alarm Root Cause Analysis. *Computer Networks,* 220, 109485.

Rodrigo, V., Chioua, M., Hagglund, T., and Hollender, M. (2016). Causal Analysis for Alarm Flood Reduction. *IFAC-PaperOnline,* 49(7), pp. 723–728.

Noroozifar, A., and Izadi, I. (2019). Root Cause Analysis of Process Faults Using Alarm Data. *27th Iranian Conference on Electrical Engineering,* 1118–1122.

Zhang, P., Hu, W., Zhang, X., and An, J. (2022). A Causal Fusion Inference Method for Industrial Alarm Root Cause Analysis Based on Process Topology and Alarm Event Data. *Journal of Beijing Institute of Technology,* 31(4), 371-381.

Venkidasalapathy, J. A., and Kravaris, C. (2022). Hidden Markov Model Based Fault Diagnoser Using Binary Alarm Signals with an Analysis on Distinguishability. *Computers & Chemical Engineering,* 160, 107689.

Venkidasalapathy, J. A., and Kravaris, C. (2020). Hidden Markov Model Based Approach for Alarm Rationalization. *21th IFAC World Congress,* 13767–13770.

Hu, W., Chen, T., and Shah, S. L. (2018). Detection of Frequent Alarm Patterns in Industrial Alarm Floods Using Itemset Mining Methods. *IEEE Transactions on Industrial Electronics,* 65(9), 7290–7300.

ANSI/ISA-18.2, 2016. ANSI/ISA-18.2: American National Standard Institute (ANSI)/International Society of Automation (ISA), Management of Alarm Systems for the Process Industries.

Cheng, Y., Izadi, I., and Chen, T. (2013). Pattern Matching of Alarm Flood Sequences by a Modified Smith–Waterman Algorithm. *Chemical Engineering Research and Design*, 91 (6), 1085–1094.

Yang, G., Hu, W., Cao, W., and Wu, M. (2020). Simulating Industrial Alarm Systems by Extending the Public Model of a Vinyl Acetate Monomer Process. *39th Chinese Control Conference,* 6093–6098.

Parvez, M. R., Hu, W., and Chen, T. (2022). Real-time Pattern Matching and Ranking for Early Prediction of Industrial Alarm Floods. *Control Engineering Practice,* 120, 105004.

Machida, Y., Ootakara, S., Seki, H., Hashimoto, Y., Kano, M. (2016). Vinyl Acetate Monomer (VAM) Plant Model: A New Benchmark Problem for Control and Operation Study. *IFAC-PaperOnline,* 49(7), 533–538.