



# An accelerated alignment method for analyzing time sequences of industrial alarm floods



Cen Guo<sup>a</sup>, Wenkai Hu<sup>b</sup>, Shiqi Lai<sup>b</sup>, Fan Yang<sup>a,\*</sup>, Tongwen Chen<sup>b</sup>

<sup>a</sup> Tsinghua National Laboratory for Information Science and Technology (TNList) and Department of Automation, Tsinghua University, Beijing 100084, PR China

<sup>b</sup> Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6G 1H9

## ARTICLE INFO

### Article history:

Received 19 August 2016

Received in revised form 24 June 2017

Accepted 26 June 2017

### Keywords:

Industrial alarm systems

Alarm floods

Sequence alignment

## ABSTRACT

In industrial processes, analyzing and predicting process faults are quite important, which could help operators to take timely and effective responses to ensure process safety and prevent further losses, especially during alarm floods. Various fault analysis methods have been proposed so far, among which the alarm flood sequence alignment (AFSA) methods, unlike other traditional model-based or statistical methods, provide fault inference from the perspective of alarm sequence similarity assessment. A new AFSA method, the match-based accelerated alignment (MAA) is proposed to generate insightful and informative alarm sequence alignments. MAA focuses mainly on alarm match analysis and outperforms other methods in terms of robustness towards nuisance alarms and improved computational efficiency. More importantly, the alarm time information is well considered and explored in MAA, rendering its alignment results capable of revealing the real similarity of alarm floods. Numerical examples and a real chemical plant case are studied to demonstrate the effectiveness and efficiency of the proposed MAA method.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

### 1.1. Background

It is widely recognized that alarm systems are of great significance to the safe operation of industrial processes. With the improvement in hardware and software design and applications, the number of alarm configurations in alarm systems increases rapidly, bringing convenience to monitoring process states and on the other hand, leading to the occurrence of alarm floods. As suggested by the industrial standard ANSI/ISA-18.2, IEC 62682, and EEMUA [1–3], an alarm flood is present if more than 10 alarms arise within a period of 10 min. When an alarm flood occurs, operators could be overwhelmed with tackling a large amount of alarms and therefore would not be able to properly handle the process abnormality in time [4,5]. Thus, there is necessity to develop techniques to deal with the problem of alarm floods and assist operators to take timely and effective actions to maintain healthy process operations.

Some research work has been carried out for the treatment of alarm floods. A method based on the dynamic fault tree was

proposed and implemented for alarm filtering and fault event dependency analysis [6]. A state-based alarming and alarm load shedding approach was used to mitigate the effects of alarm floods [7]. Plant connectivity and alarm logs were combined to reduce the number of alerts in an automation system [8]. A dynamic alarm management strategy was provided to deal with alarm floods that occur during the transition of operating states by changing alarm limits according to a priori transition knowledge [9]. In order to remove chattering and repeating alarms, an online method based on alarm duration and interval analysis was proposed [10]. And for the same purpose, a quantitative measure employing statistical property of the process variables and alarm parameters was put forward [11]. The visualization of alarm floods has been studied as well. A new visualization method was employed to facilitate alarm presentation by showing alarms in time series together with short alarm descriptions [12]. A technique using principal components in parallel coordinates was implemented for process monitoring applications where a large number of sensors are used to detect and identify abnormal operation conditions [13]. In addition, methods to assist operators during industrial alarm floods based on analyzing historical alarm flood data were proposed to improve alarm system performance [14,15]. Moreover, alarm flood sequence alignment (AFSA) is an interesting way of analyzing alarm floods based mainly on alarm sequence similarity assessment.

\* Corresponding author.

E-mail address: [yangfan@tsinghua.edu.cn](mailto:yangfan@tsinghua.edu.cn) (F. Yang).

The agglomerative hierarchical clustering (AHC) method and the dynamic time warping (DTW) technique were employed for alarm flood clustering and alignment [16]. A strategy specially designed to deal with the time cross phenomenon in the alignment based on the Smith–Waterman algorithm was proposed [17]; it was further developed for multiple sequence pattern mining [18]. The modified basic local alignment search tool (BLAST) was put forward to reduce the search space of the alignment process [19]. The match-based accelerated alignment method (MAA) proposed in this paper is an AFSA method.

## 1.2. Basics of AFSA

The attributes of alarms typically include tag name, alarm identifier, time stamp, alarm priority, unit area, and process description [20]. Tag name and alarm identifier compose the alarm tag, e.g., tag name 001 and identifier PVHI (process variable high alarm) compose alarm tag 001.PVHI. One alarm tag and one time stamp make up the alarm tuple used in MAA as

$$x = (e, t), \quad (1)$$

where  $x$  represents an alarm,  $e$  denotes the alarm tag, and  $t$  denotes the alarm time. For a given alarm flood sequence

$$\begin{aligned} X &= \{x_1, x_2, \dots, x_{N_X}\} \\ &= \{(e_{x_1}, t_{x_1}), (e_{x_2}, t_{x_2}), \dots, (e_{x_{N_X}}, t_{x_{N_X}})\}, \end{aligned} \quad (2)$$

the numerical alphabet set for alarm tags is expressed as

$$\Sigma = \{1, 2, \dots, N_\Sigma\}, \quad (3)$$

and therefore  $e_{x_i} \in \Sigma, i = 1, 2, \dots, N_X$ .

The alignment of two alarm flood sequences is typically composed of a series of matches, mismatches, and gap-matches. The three basic alignment components are defined as follows. For alarm  $x_i$  in sequence  $X = \{x_1, x_2, \dots, x_{N_X}\}$  and alarm  $y_j$  in sequence  $Y = \{y_1, y_2, \dots, y_{N_Y}\}$ , if  $e_{x_i} = e_{y_j}$ , the two alarms form a match, expressed as

$$\langle x_i, y_j \rangle \text{ or } \langle (e_{x_i}, t_{x_i}), (e_{y_j}, t_{y_j}) \rangle. \quad (4)$$

If  $e_{x_i} \neq e_{y_j}$ ,  $x_i$  and  $y_j$  form a mismatch. And if no proper  $y_j$  can be aligned to  $x_i$ ,  $x_i$  and a gap ‘-’ form a gap-match, expressed as

$$\langle x_i, - \rangle \text{ or } \langle (e_{x_i}, t_{x_i}), - \rangle. \quad (5)$$

Similarity scores are assigned to the matches, mismatches, and gap-matches in the alignment. The similarity score for a match is usually positive, while the scores for a mismatch and a gap-match are always negative or zero. Thus, the mismatch and the gap-match scores are also called the mismatch penalty and the gap-match penalty.

The objective of sequence alignment is to find a way of aligning two sequences that corresponds to a high overall similarity score. Dynamic programming is commonly employed for the alignment realization. Procedures of the dynamic programming algorithm used by a local alignment method, the Smith–Waterman method, to align sequences  $X$  and  $Y$  are described as follows.

1. Build an  $(n + 1)$ -by- $(m + 1)$  dimensional similarity score matrix  $\mathbf{H}$  and initialize its first row and column to be 0.
2. Calculate the elements in  $\mathbf{H}$  from the upper left corner to the lower right corner by

$$H_{i+1,j+1} = \max\{H_{i,j} + f(x_i, y_j), H_{i,j+1} + \delta, H_{i+1,j} + \delta, 0\}, \quad (6)$$

where

$$f(x_i, y_j) = \begin{cases} 1, & \text{if } e_{x_i} = e_{y_j}, \\ \mu, & \text{if } e_{x_i} \neq e_{y_j}. \end{cases} \quad (7)$$

Here  $\mu$  is the mismatch penalty and  $\delta$  is the gap-match penalty. The default match score equals 1 in this work.

- Find the maximum element value in  $\mathbf{H}$  which is the similarity score of the alignment of  $X$  and  $Y$ .
- Trace the way of the generation of the maximum value backwards until meeting a value of 0, and the obtained path corresponds to the optimal local alignment.

Moreover, alarms triggered by physically connected devices usually arise simultaneously and their chronological order is of little importance as long as the alarm occurrence time is close. For example, considering alarms  $x_1, x_2, y_1$ , and  $y_2$  ( $e_{x_1} = e_{y_1}$  and  $e_{x_2} = e_{y_2}$ ), if  $t_{x_1} > t_{x_2}$  and  $t_{y_1} < t_{y_2}$  but  $t_{x_1}$  is approximately equal to  $t_{x_2}$ , the four alarms can form matches  $\langle x_1, y_1 \rangle$  and  $\langle x_2, y_2 \rangle$  even though there is time cross among the alarms. Conventional sequence alignment methods are strict with the alarm order and prohibit any time cross to be present in the alignment. In view of this, the work in [17] put forward a strategy to enhance the tolerance of alarm time cross for AFSA and proposed the modified Smith–Waterman (MSW) method to align alarm flood sequences of poor data quality. The time cross treatment strategy is adopted in MAA as well and thus is introduced briefly as below.

Define the time distance vector  $d_i$  for alarm  $x_i$  in sequence  $X$  as

$$d_i = [d_{i1} \ d_{i2} \ \dots \ d_{iN_\Sigma}]^T, \quad (8)$$

where

$$d_{ik} = \min_{1 \leq j \leq N_X} \{|t_{x_i} - t_{x_j}| : e_{x_j} = k\}. \quad (9)$$

Here  $k$  is the numerical symbol in the alarm tag alphabet  $\Sigma$ . Each  $d_{ik}$  indicates the time span between alarm  $x_i$  and the closest alarm with alarm tag  $k$ . Subsequently, we have the time weighting vector  $w_i$  for  $x_i$  as

$$\mathbf{w}_i = [w_{i1} \ w_{i2} \ \dots \ w_{iN_\Sigma}]^T, \quad (10)$$

where

$$w_{ik} = e^{-\frac{d_{ik}^2}{2\sigma^2}}. \quad (11)$$

The standard deviation  $\sigma$  is selected according to how much weight the user wants to put on the close-by alarms to blur their orders in the alignment. Finally, the similarity score in Eq. (7) is reformulated as

$$f(x_i, y_j) = \mu + (1 - \mu) \max_{1 \leq k \leq N_\Sigma} w_{ik}^x w_{jk}^y + \epsilon, \quad (12)$$

where  $w_{ik}^x$  and  $w_{jk}^y$  denote the time weights for alarm  $x_i$  in  $X$  and alarm  $y_j$  in  $Y$ , and  $\epsilon$  is a negative constant to adjust the score if  $e_x \neq e_y$ ; for cases where  $e_x = e_y$ , set  $\epsilon$  to be 0.

## 1.3. Motivations

Existing AFSA methods suffer from some common application problems in practical scenarios. First, the existence of nuisance alarms could have significant influence on the alignment results; second, the computational burden of dynamic programming is always high, especially when the considered sequences are combined by a large number of alarms; and last but not the least, time information of alarms is not appropriately incorporated into the

Download English Version:

<https://daneshyari.com/en/article/4998410>

Download Persian Version:

<https://daneshyari.com/article/4998410>

[Daneshyari.com](https://daneshyari.com)