Journal of Process Control 37 (2016) 21–33

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

A study of polynomial fit-based methods for qualitative trend analysis

Bo Zhou, Hao Ye*

Department of Automation, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 14 April 2015
Received in revised form 6 October 2015
Accepted 7 November 2015
Available online 29 November 2015

Keywords:
Qualitative trend analysis
Trend extraction
Polynomial fit
Process monitoring and diagnosis

ABSTRACT

Qualitative trend analysis (QTA) of sensor data is a useful tool for process monitoring, fault diagnosis and data mining. However, because of the varying background noise characteristics and different scales of sensor trends, automated and reliable trend extraction remains a challenge for trend-based analysis systems. In this paper, several new polynomial fit-based trend extraction algorithms are first developed, which determine the parameters automatically in the hypothesis testing framework. An existing trend analysis method developed by Dash et al. (2004) is then modified and added to the abovementioned trend extraction algorithms, which form a complete solution for QTA. The performance comparison of these algorithms is made on a set of simulated data and Tennessee Eastman process data based on several metrics.

1. Introduction

Qualitative trend analysis (QTA) has proven to be an efficient tool for extraction and interpretation of useful high-level knowledge from large volumes of process data and has been successfully applied in the fields of process monitoring and fault diagnosis [1–5]. However, it would be an overload for operators to monitor the trends in huge sets of sensor measurements on a regular basis. Thus, automated and accurate trend extraction has great potential in both the process monitoring and fault diagnosis of complex chemical processes and medical systems [6]. The concise form of qualitative trend analysis enables it to capture the significant events happening in the processes and has also made it a useful tool in the data compression and data mining fields [7].

A trend is a consecutive sequence of episodes in which the beginning and end time instants are uniquely determined to manifest a time of qualitative state change [1–5]. In QTA, an alphabetic symbol is used to characterize the temporal behaviors during each episode, which is called primitive [8,9]. In the 1990s, Cheung and Stephanopoulos [10] built a formal methodology to represent the process trends using the so-called triangular episode. This temporal episode is characterized by a triangulation that consists of an initial slope, a final slope and an average slope connecting the boundary time points. Janusz and Venkatasubramanian [11] developed a qualitative trend description language whose fundamental elements are primitives. In their work, seven primitives with constant signs of first and second derivatives are used to characterize the important qualitative information in the process trends, which are depicted in Fig. 1. To describe more complex trend behaviors, many other researchers have expanded the set of primitives to incorporate discontinuous or even unknown derivatives. Interested readers can refer to [12,13] for further details. In this paper, the seven primitives defined by [11] are adopted.

The main tasks of QTA include (1) extracting the trends—i.e., segmenting the signal into non-overlapping episodes—and (2) analyzing the trends—i.e., assigning primitives to each of the episodes and designing a map from primitive sequences to process states [13,14]. Among the existing QTA methods, some focus on task (1), such as [15–18,22], and many other papers execute both task (1) and task (2), including [1,3–5,9,13,14,19–21].

Owing to the various scales of the underlying driving events happening in the process, the variables in the industrial process usually have different intrinsic dynamics and are often corrupted by different levels of noise, which complicates the trend identification task [13,14]. Existing methods for task (1) (i.e., trend extraction) can be mainly classified into polynomial fit-based methods and basis function projection-based methods [8]. Owing to its shorter computational time and higher robustness to noise, the polynomial fit-based method has gained popularity through many previous studies [22]. Keogh et al. [17] further grouped the polynomial fit-based trend extraction algorithms into three classes—namely, sliding window-, top-down-, and bottom-up-based methods. Later, the fixed-width window approach was also added as the fourth type of polynomial fit-based algorithm in the literature [8].

The first class of polynomial fit-based trend extraction algorithms—i.e., SWTE (sliding window trend extraction)

* Corresponding author. Tel.: +86 010 62790497; fax: +86 010 62786911.
E-mail address: haoye@tsinghua.edu.cn (H. Ye).

http://dx.doi.org/10.1016/j.jprocont.2015.11.003
0959-1524/

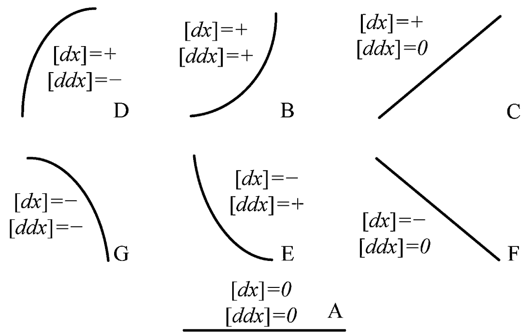

Fig. 1. Seven primitives.

algorithms—add the data points to the current window until the fitting error exceeds a predefined threshold and start a new segment when the current approximation is no longer acceptable [17]. Among the SWTE algorithms, the piecewise linear online trending (PLOT) algorithm proposed in the literature [18] chooses the statistical prediction interval as the threshold. In [6,15,20,23], Charbonnier and Portet used the cumulative sum of the difference between the online data and the model output as another form of fitting error. The threshold for the cumulative sum must be tuned for each signal separately in [6,20]. In [23], the threshold for the cumulative error is allowed to switch between two sets of values according to the variance of the signal rather than remain fixed. In [15], the threshold for the cumulative error is self-tuned online based on the estimation of the noise level with a median filter, which considers the signal characteristics adaptively. The second class of polynomial fit-based trend extraction algorithms—i.e., TDTE (top-down trend extraction) algorithms—first fit the entire data set with one linear model and split the dataset into two sets if the fitting error is larger than a given threshold. This splitting procedure proceeds in the subsequences until the fitting error is below the threshold [17]. The splitting point and the stopping criteria should be selected based on some critical parameters in the execution of TDTE. Examples can be found in [9,16,24]. The third class of polynomial fit-based trend extraction algorithms—i.e., BUTE (bottom-up trend extraction) algorithms—work by merging the sequences from the finest approximations ($n/2$ segments for the signal of length n). The pair of adjacent segments with the minimal merging cost is selected to merge at each iteration. This greedy merging process is repeated until the minimal merging cost exceeds the predefined threshold [17]. In [26–28], the fitting error is chosen as the merging cost, and the threshold is specified by the user. The fourth class of polynomial fit-based trend extraction algorithms—i.e., FWTE (fixed-width window trend extraction) algorithms—divide the dataset into segments of equal length, which is the simplest method of signal segmentation [8].

In all four types of existing polynomial fit-based trend extraction algorithms, the selection of parameters (e.g., thresholds, noise level) plays an important role. In most of the abovementioned existing methods, the parameters are either constant values specified by the user or are tuned for each signal heuristically; therefore, the selection lacks a statistical basis. An exceptional work is found with the techniques proposed by Dash et al. [9]. In their work, both the significance of fitting error with respect to noise and the significance of the derivatives are examined by hypothesis testing.

In addition to trend extraction, Dash et al. [9] further propose a trend analysis algorithm, which belongs to task (2) of QTA mentioned above. In their method, the signs of the first second derivatives are determined by the significance test of the derivatives using the t-test. The assignment of primitives can then be made based on the signs of the derivatives. However, two limitations exist in their work: (1) there exists an approximation in the calculation of the sign of the first derivative; (2) some possible cases are not considered in the assignment of primitives (please refer to Section 3 for details).

The main goals of this paper can be summarized as follows. (1) Because most of the polynomial fit-based trend extraction algorithms [6,16,20,23–27] for task (1) of QTA require manual tuning of the parameters (except for the method of Dash et al. [9], which belongs to TDTE), in this paper, the idea of automatically determining parameters in the framework of hypothesis testing proposed by Dash et al. [9] is extended to the SWTE and BUTE algorithms. (2) Considering the abovementioned limitations of Dash's algorithm for trend analysis (i.e., task (2) of QTA) [9], some modifications are presented in this paper; this is called the modified algorithm for trend analysis for convenience. (3) Except for the work of Dash et al. [9], most algorithms in [6,16,20,23–27], which are SWTE or BUTE algorithms, only address the issue of trend extraction (i.e., task (1) of QTA) and do not carry out trend analysis (i.e., task (2) of QTA). Thus, in this paper, the modified trend analysis method given in (2) is further added to the SWTE and BUTE algorithms for trend extraction, which results in a complete solution for QTA. (4) A comprehensive comparison of the different QTA algorithms is made; this has not been given in the abovementioned literature.

The remainder of the paper is organized as follows. Section 2 briefly reviews the methods proposed by Dash et al. for trend extraction (i.e., task (1) of QTA) and trend analysis (i.e., task (2) of QTA). Section 3 presents the modifications to Dash's algorithm for trend analysis. In Section 4, the automatic method of determining parameters in the framework of hypothesis testing is extended to SWTE and BUTE algorithms. The modified trend analysis algorithm is then added to SWTE and BUTE algorithms. In Section 5, a quantitative comparison of the different QTA algorithms is drawn using simulated data and Tennessee Eastman process datasets. Section 6 finishes the paper with discussions and conclusions.

2. A brief introduction to Dash's original framework

Dash's method mainly consists of two steps corresponding to tasks (1) and (2) of QTA: (1) extract the trend through a polynomial fit-based interval-halving technique; (2) analyze the trend by assigning primitives to the segments based on the signs of the derivatives [9]. In the following, these two steps will be explained separately.

2.1. Step 1: trend extraction

In [9], a polynomial $\hat{y}_i = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \cdots + \beta_m t_i^m$ is used to fit the process signal y_i, $i = 1, \ldots, n$. Without loss of generality, the time window is normalized to [0,1]. Standard least square estimation of the coefficients $\hat{\boldsymbol{\beta}} = [\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_m]^T$ is adopted; i.e.,

$$\hat{\boldsymbol{\beta}} = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{Y} \tag{1}$$

with $(\mathbf{T})_{j,k} = t_j^{k-1}$, $1 \le j \le n$, $1 \le k \le m$, $Y = [y_1, y_2, \ldots, y_n]^{\mathrm{T}}$ and the following fitting error

$$\varepsilon_{fit}^2 = \frac{1}{v_1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

where $v_1 = n - m - 1$ denotes the degree of freedom in the estimation of ε_{fit}^2.

In [9], a constant polynomial with order $m = 0$ is fitted first. If the fitting error is statistically significant compared to the noise level, then the linear fit with order $m = 1$ is attempted. If even the quadratic polynomial with order $m = 2$ cannot approximate the data with acceptable fitting error, the window length is interval-halved, and the process is iteratively repeated on the first-half data until