Contents lists available at ScienceDirect

Chemical Engineering Research and Design

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

Adaptive soft sensor based on time difference Gaussian process regression with local time-delay reconstruction

Weili Xiong[a,b], Yanjun Li[b], Yujia Zhao[c], Biao Huang[c,*]

[a] Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, China
[b] School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China
[c] Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta T6G2G6, Canada

ARTICLE INFO

Article history:
Received 8 September 2016
Received in revised form 14 November 2016
Accepted 20 November 2016
Available online 25 November 2016

Keywords:
Soft sensor
Time difference model
Moving window
Time-delay extraction
Dataset reconstruction

ABSTRACT

Apart from strong nonlinearity and time-varying behaviors in industrial processes, the hidden time-delay information, which is unfortunately overlooked in most existing modeling methods, should also be taken into account in soft sensor modeling. In view of this, a novel soft sensor, referred to as local time-delay reconstruction based moving window time difference Gaussian process regression (LTR-MWTDGPR), is proposed in this paper. To deal with the time-delay, a fuzzy curve analysis based local time-delay parameter extraction procedure is performed along with a strategy of a moving window, which simultaneously captures the process time-varying feature. Then the local window training dataset and new query sample are reconstructed according to the time-delay parameter set at the next sampling instant. Afterwards, the time difference Gaussian process regression is employed to handle the drifting feature of local reconstructed dataset. The effectiveness and accuracy of the proposed LTR-MWTDGPR approach in predicting quality variables are verified through a real sulfur recovery unit and an industrial debutanizer column.

1. Introduction

There is a great demand on control and optimization of product quality in modern industrial processes, which leads to requirements of the online measurement of process variables (Fortuna et al., 2007). In many practical applications, quality-related variables such as concentration in gas flow and certain chemical ingredients of products are difficult to measure online (Ahmad et al., 2014; Yan et al., 2004). In such circumstance, soft sensors have been extensively applied through constructing mathematical models between auxiliary process variables and the dominant quality variables (Facco et al., 2009; Khatibisepehr et al., 2013).

In some cases, although there are online analyzers installed for quality variables on site, the measurement sequence of quality variables is not consistent with the sampling sequence of process variables, exhibiting significant time-delay introduced by signal and material transmission, or installation location and analyzing cycle of measuring instruments (Fortuna et al., 2007). If we ignore such time-delay, the modeling accuracy and control quality of system would be greatly compromised. Increased delay would come along with deteriorated control performance. Therefore, it is imperative to have a reliable estimate of the delay between process variables and quality variables to optimize the control of chemical production processes.

As process time-delay plays a critical role in system dynamics and control, there are numerous published works on the identification of time delay systems (Richard, 2003; Bozorg and Davison, 2006; Tufa and Ramasamy, 2011), and the topic on how to develop reliable online soft sensors in presence of time delays has attracted much attention. In order to extract the process delay information, Fortuna et al. (2005) made use of the designing parameters of process hardware instruments, such as reactor volume or the length of a pipe, to estimate the approximate delay range of the device, using a nonlinear autoregressive moving average model structure which involves certain amount of lagged samples to overcome the large time-delay introduced by gas chromatograph; besides, by computing the correlation coefficient

* *Corresponding author*. Fax: +1 780 492 2881.
E-mail address: bhuang@ualberta.ca (B. Huang).
http://dx.doi.org/10.1016/j.cherd.2016.11.020

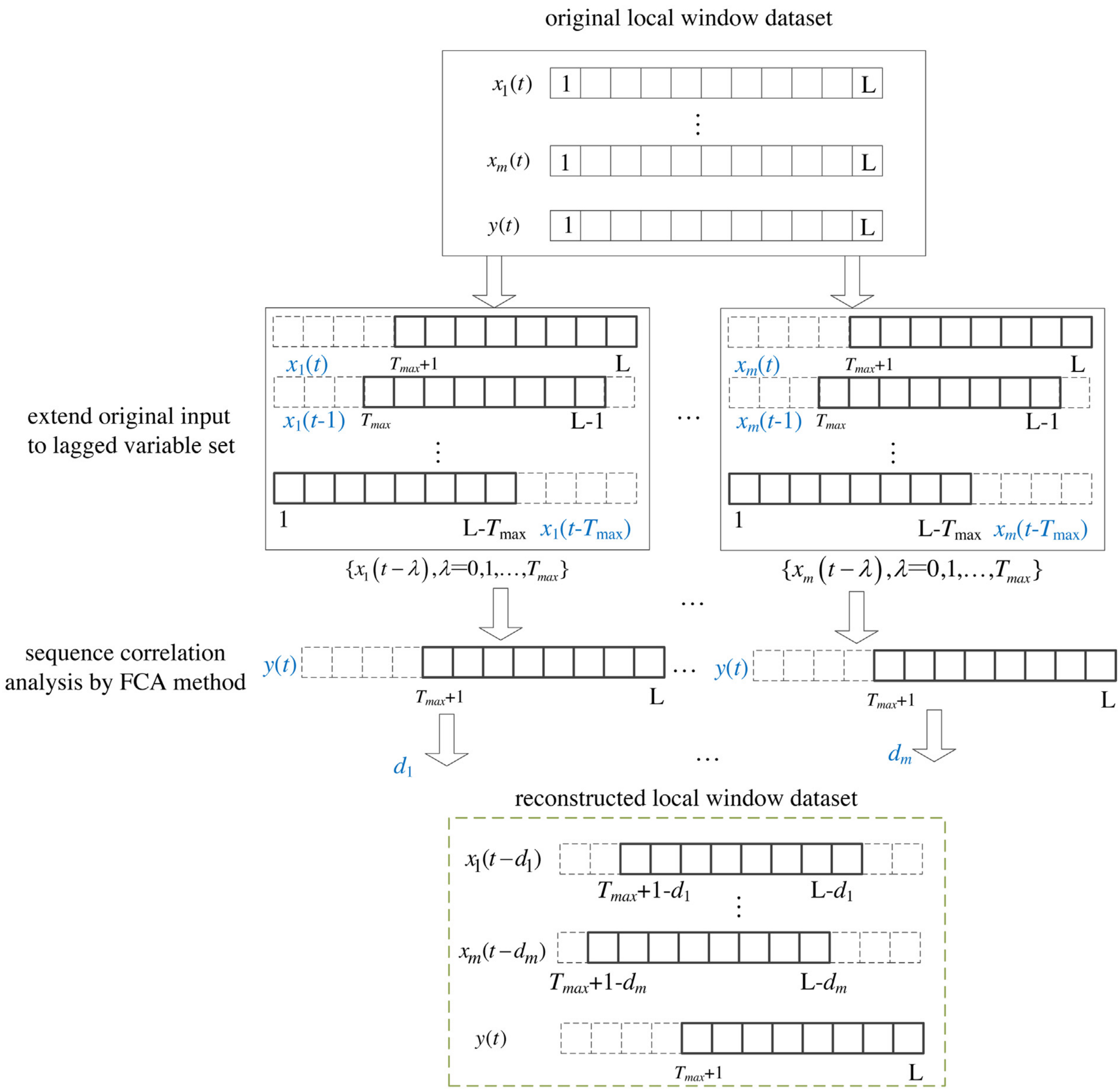

Fig. 1 – The flow diagram of FCA-LDR procedure.

between input and output variables, Komulainen et al. (2004) and Zhang et al., (2006) constructed an online dynamic partial least squares model and an aggregated neural network respectively to refine soft sensor models with lagged sample information. In addition, Souza et al., (2010) developed an artificial neural network based data-driven soft sensor, utilizing mutual information index of process and quality variables to introduce variable delay into soft sensor model and select informative input variables to further improve the soft sensor reliability. However, there are still a number of open problems unsolved. For example, the number of lagged samples to be used in ARMA model structure is obtained by trial and error, which is prone to unstable model performance. Correlation coefficient based delay estimation methods are limited to linear systems, and mutual information based algorithms tend to show a high degree of computational complexity and need a lot of data. Although sometimes the delay parameters can be determined through prior knowledge from in-depth analysis of process mechanism, such an estimation procedure might exhibit characteristics of randomness and uncertainty. Thus, a delay estimation method that can well describe process nonlinearity with a relatively low computational load is much needed. In the same time, since the operating conditions of process control are often time-varying, process data presents clear stage-wise characteristics. Therefore, when making an estimate of delay parameters, time-delay and shifting features under different operating conditions should be both considered so as to better capture local characteristics and provide a more reliable soft sensor model for process quality control.

Currently, data-driven methods have been widely applied in soft sensor modeling due to the fact that they do not need much prior process knowledge and are simple to develop compared with the first principle models (Yuan et al., 2016; He et al., 2016; Gholami et al., 2015; Kadlec et al., 2009). Data-driven models, just as the name implies, are built on numerous data collected from the process historical database. Although synchronous data acquisition can be done in large scale with the rapid development of distributed control systems, the time-delay between process variables and quality variables still exists due to the different spatial and temporal distribution of process instruments. Thus the real-time collected dataset contains useful delay information, which provides opportunities for the establishment of time-delay related soft sensor models. Given that delay is almost ubiquitous in industrial processes, active steps must be the taken to correct dataset by re-matching input and output time sequence, which will bring a lot of benefits to subsequent model building, real-time quality estimation, and online correction.

On the other hand, nonlinearity and time-varying behavior of chemical industrial processes are two main topics to discuss throughout soft sensor development history. When a soft sensor model is established and put into service, the problem of model degradation is inevitable, which can lead to deterioration in prediction accuracy of