Computers and Chemical Engineering 69 (2014) 66–74

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Simultaneous data reconciliation and gross error detection for dynamic systems using particle filter and measurement test

Zhengjiang Zhang[a,b], Junghui Chen[b,*]

[a] College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou 325035, People's Republic of China
[b] Department of Chemical Engineering, Chung Yuan Christian University, Chungli, Taiwan 320, Republic of China

ARTICLE INFO

Article history:
Received 5 February 2014
Received in revised form 18 June 2014
Accepted 27 June 2014
Available online 7 July 2014

Keywords:
Chemical processes
Data reconciliation
Gross error detection
Measurement test
Particle filter
System engineering

ABSTRACT

Good dynamic model estimation plays an important role for both feedforward and feedback control, fault detection, and system optimization. Attempts to successfully implement model estimators are often hindered by severe process nonlinearities, complicated state constraints, systematic modeling errors, unmeasurable perturbations, and irregular measurements with possibly abnormal behaviors. Thus, simultaneous data reconciliation and gross error detection (DRGED) for dynamic systems are fundamental and important. In this research, a novel particle filter (PF) algorithm based on the measurement test (MT) is used to solve the dynamic DRGED problem, called PFMT-DRGED. This strategy can effectively solve the DRGED problem through measurements that contain gross errors in the nonlinear dynamic process systems. The performance of PFMT-DRGED is demonstrated through the results of two statistical performance indices in a classical nonlinear dynamic system. The effectiveness of the proposed PFMT-DRGED applied to a nonlinear dynamic system and a large scale polymerization process is illustrated.

1. Introduction

To control, optimize, or evaluate the behavior of a chemical process, it is important to know the current status of the process. Since it is generally difficult to measure all the system state variables, the values of the process variables contained in the model are chosen to represent the operation plant. The dynamic data reconciliation strategy uses the state space model and partial process measurements to estimate unmeasured state variables and reconcile the measurements simultaneously. However, process measurements generally contain noise and gross errors, both of which have a detrimental effect on process control and optimization which require accurate estimated states and measurements. When measurements are corrupted by random variations of the environment, they are said to be affected by noise. Gross errors in measurements usually occur for many different reasons. These include human errors, instrumental errors, fraudulent behavior, and faults in systems. The presence of gross errors affects the results of dynamic data reconciliation since the large errors are not sufficiently eliminated or corrected. As a result of smearing, both the reconciled measurements and the estimates of states may become less accurate. Clearly, the measurements with gross errors need to be identified and the magnitudes of the gross errors should be estimated to decrease their impact on the results of data reconciliation. Gross error detection and data reconciliation are generally considered crucial techniques. Therefore, simultaneous data reconciliation and gross error detection (DRGED) for dynamic systems are fundamental and important to process control and optimization.

Data reconciliation is also a model-based filtering technique that attempts to reduce the inconsistency between measured process variables and a process model. In the study of the dynamic linear system, the Kalman Filter (KF) plays a crucial role in measuring data with random errors (Kalman, 1960; Kalman and Bucy, 1961). KF estimates the desirable statistical properties of being unbiased and also has the minimum variance under the assumption of the Gaussian distribution. Furthermore, KF estimates are the maximum-likelihood estimates (Sage and Melsa, 1971). In practice, many physical systems exhibit nonlinear dynamics and have states subject to hard constraints such as non-negative concentrations and temperatures. Therefore, KF (designed for unconstrained linear systems) is no longer directly applicable. Jazwinski (2007) extended the use of KF to nonlinear dynamic systems, and proposed the extended KF (EKF). Stanley and Mah (1977) were the first ones to tackle the data reconciliation process using EKF (Narasimhan and Jordache, 2000). Since the reliability of EKF-based approaches often decreases as the system nonlinear complexities and

* Corresponding author. Tel.: +886 3 2654107; fax: +886 3 2654199.
E-mail address: jason@wavenet.cycu.edu.tw (J. Chen).

http://dx.doi.org/10.1016/j.compchemeng.2014.06.014
0098-1354/

modeling uncertainties increase, the model should not be complex. A simple random walk model can be used to characterize the process dynamics. Later, Robertson et al. (1996) presented a formulation of the dynamic data reconciliation problem as a special case of a more general moving-horizon state estimation method. In order to detect the gross error, Singhal and Seborg (2000) proposed a probabilistic formulation that combined the EKF and the expectation-maximization (EM) algorithm in measurement reconciliation. Another improved KF filter, the unscented KF (Julier et al., 2000; Romanenko and Castro, 2004; Qu and Hahn, 2009), was developed for the application of KF. Due to linearization at each time step for the EKF application, large errors and divergence of the filter may occur (Romanenko et al., 2004). Furthermore, if the state and/or measurement equations are highly nonlinear and the posterior distribution of the states is non-Gaussian, the filters will give unsatisfactory state estimates in a number of applications (Chen et al., 2005, 2008).

For the optimal solution to the nonlinear filtering problem, a complete description of the conditional probability density should be maintained and is infinite dimensional (Kushner, 1967). During the past decade, particle filtering (PF) technique has become a popular signal processing tool for problems that involve nonlinear tracking of an unobserved signal of interest given a series of related observations (Doucet et al., 2001; Arulampalam et al., 2002; Maiz et al., 2012; Gning et al., 2012). Compared with KF, PF is described as a general filter for the nonlinear and non-Gaussian state space systems. It does not assume a fixed shape of any probability density. Instead, it approximates the densities of interest via the Monte Carlo estimation method to generate a large number of random samples (particles) for constructing the approximations of the posterior probability distribution of the states. PF can capture the time-varying nature of distributions commonly encountered in nonlinear dynamic problems. This can be computed from the sampled particles at any moment. Some papers, but not many, applied to PF in data reconciliation have been reported in literature. Chen and his coworkers used PF for dynamic data reconciliation and process change detection (Chen et al., 2008). They also applied PF and the kernel smoothing method on-line for state and parameter estimation in a highly non-linear batch processes (Chen et al., 2005). López-Negrete et al. (2011) used the constrained PF approach to approximate the arrival cost in moving horizon estimation and applied the method in the continuously stirred tank reactor and the constrained batch reactor process in order to estimate the unmeasured states accurately. Zhao et al. (2013) investigated a parameter estimation problem for batch processes using EM algorithm with PF.

The performance of PF may be severely degraded when gross errors are present in the process measurements. In only a few papers, (Chen et al., 2008) the dynamic reconciliation of measurements with gross errors is considered in the dynamic process system. The measurements with gross errors should not be treated in the same way as the regular measurements. The detection and processing of gross errors should be combined in the procedure of particle filtering. Chen et al. (2008) proposed a method of the dynamic data reconciliation using PF for Gaussian distribution based gross error detection. The goal is to improve performance by detecting gross errors rather than processing them. If the gross error is present in some measurements, the reconciled data using PF may be inaccurate because PF procedures would be deteriorated by a very small number of highly weighted particles amongst a hoard of almost useless particles carrying a tiny proportion of the probability mass. Those particles result in failure because of an inadequate representation of the required probability density function (PDF). It is, therefore, necessary to identify and remove such gross errors. In this paper, a new PF scheme, the PF and MT (measurement test) based simultaneous Data reconciliation and gross error detection method (PFMT-DRGED), is proposed. In most literature, several data processing methods based on some particular simulation runs were shown to have multiple advantages. Since PF is a statistical method, it is unfair to use some particular simulation runs as a comparison. Statistical performance indices should be used to evaluate the performance of PF. Therefore, two statistical performance indices, the observed power (OP) and the average number of Type I errors (AVTI), are used here to evaluate the performance of this method.

The rest of the paper is organized as follows. The procedures of PF for the dynamic data reconciliation are described in the next section. MT for gross error detection is then introduced. The integrated implementations of PFMT-DRGED are proposed in Section 3. In Section 4, the effectiveness of PFMT-DRGED is demonstrated through the DRGED problems in two case studies. The first case is simple. It is widely used in the dynamic data reconciliation problems. It can make fair comparisons under the same basis. The second case which discusses the free radical polymerization of styrene, illustrates how PF can be applied to a realistic system with the large-scale and nonlinear processes. Finally, conclusions are drawn in Section 5.

2. PF for data reconciliation problem in nonlinear dynamics

State estimations deal with the problem of inferring knowledge about process variables (or state) indirectly measured from possibly noisy observations in a real process, and the state is a physical quantity that affects the observation represented by a certain process model (Shao et al., 2009). With the Bayesian view on the estimation, both the states and the observations are stochastic entities. Thus, the inference result is a conditional density function of the states given the observational outcomes. PF provides a robust tracking framework as it models uncertainty. It can consider multiple state hypotheses simultaneously. Since states are less likely to temporarily remain in the tracking process, PF can deal with short-lived occlusions. This section describes how to apply PF to the data reconciliation problems in the nonlinear dynamics.

A typical model for the dynamic system contains the state dynamics and the measurement equations,

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{v}_{k-1} \\ \mathbf{z}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{n}_k \end{aligned} \tag{1}$$

where $\mathbf{x}_k$ is the process state to be estimated and k denotes the time step. $\mathbf{f}$ and $\mathbf{h}$ are the known non-linear functions and the known measurement functions, respectively. $\mathbf{z}_k$ and $\mathbf{u}_k$ are the vector of the received measurements and the input vector at time step k, respectively. v_{k-1} and $\mathbf{n}_k$ are the white noise sequences for the process states and measurements, respectively. They are generally assumed to follow Gaussian distribution, i.e. $\mathbf{v}_{k-1} \sim G(\mathbf{v}_{k-1}\,; \mathbf{0}, \mathbf{\Psi})$, and $\mathbf{n}_k \sim G(\mathbf{n}_k\,; \mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{\Psi}$ and $\mathbf{\Sigma}$ are the covariance matrices.

The basic idea of PF is a direct mechanization of the formal Bayesian filter. Suppose that a set of random samples (also called particles) $(\{\mathbf{x}_k^i, i = 1, \ldots, N\})$ from the posterior PDF $p(\mathbf{x}_{k-1}|\mathbf{z}_{k-1})$ $(k > 0)$ are available.

The particles $\mathbf{x}_k^i$ are usually drawn from $p(\mathbf{x}_k|\mathbf{x}_{k-1}^i)$. Generally, in the prediction phase, each of these particles from time step $k-1$ in the state model (Eq. (1)) generate a set of prior samples at time step k,

$$\mathbf{x}_k^i = \mathbf{f}(\mathbf{x}_{k-1}^i, \mathbf{u}_{k-1}) + \mathbf{v}_{k-1}^i \tag{2}$$

where $\mathbf{v}_{k-1}^i$ is an independent sample drawn from PDF of the system with noise.

In the update phase, using the prior samples in the light of measurement $\mathbf{z}_k$, a weight $\{\mathbf{w}_k^i, i = 1, \ldots, N\}$ is calculated for each particle. This weight is the measurement likelihood evaluated at the value of the prior sample $w_k^i = p(\mathbf{z}_k|\mathbf{x}_k^i)$. The associated weights