242e Distributed Partially Adaptive Data Reconciliation with Intelligent Sensor Network

Yen Yen Joe, Zhong Qiang Ding, Keck Voon Ling, and Jose A. Romagnoli

Partially adaptive data reconciliation utilizes the partially adaptive estimation scheme, where the parameters of the estimators are adapted to data. The advantage of partially adaptive data reconciliation is its potential to achieve higher efficiency, as the estimator is optimized according to the data. The actual data characteristics cannot be predicted with absolute certainty, and may change over time. This limits the efficiency of non-adaptive, parametric estimators using a priori estimator parameters, as these parameters are usually determined using limited a priori knowledge of the data. The partially adaptive estimation, on the other hand, is an a posteriori estimation. Because the knowledge incorporated to the estimator is extracted from the data themselves, more accurate estimation is expected.

A trade-off for more accurate, posteriori knowledge is the extra computation burden and processing time associated with the adaptation of the estimator parameters to the data, which is not required for non-adaptive estimation. The adaptation may be performed by simultaneously estimating the estimator parameters and the process variables. However, this can result in a very large optimization problem, as for each process variable, there are a few estimator parameters to be estimated in addition to the process variable itself. An alternative is to split the estimation of estimator parameters with that of the process variables. The result is a two-level strategy: firstly, the estimation of the estimator parameters from the data; and secondly, the reconciliation of the data. We term them as the adaptation step and the reconciliation step, respectively. A reasonable assumption is that the measurement noise or data characteristics of each sensor are independent from those of the other sensors. As such, the adaptation step can be performed for each sensor, or each process variable, independently. To summarize, the steps involved in this scheme are as follows: first, for each process variable, estimate the parameters of the estimator from the data; and second, use the adapted estimator parameters to reconcile the data for all process variables.

The two-level partially adaptive strategy not only reduces the optimization size, but also facilitates further improvement in efficiency by means of distributed processing. We have proposed an intelligent sensor network framework [1], whereby the capability to process data and to communicate is incorporated to each sensor. Two-tier federated network architecture is adopted. The architecture is a network of clusters. A cluster is made up of one higher level node and a few lower level ones. The two tiers correspond to two data/ information processing levels: the lower tier is at the physical sensor level, while the higher, at the cluster level. The definition and formation of clusters is determined by how a task is distributed to the various nodes in the network. Some feasible distribution schemes include decomposing a process plant into functional sections or into unit operations. The sensors within a functional section or a unit operation are then grouped into the same cluster, with one higher level node (cluster head) to perform multivariate data processing for the cluster.

From the point of view of the two-level partially adaptive data reconciliation strategy, the distributed processing of the intelligent sensor network can be leveraged by delegating the first step, i.e. adaptation of estimator to data, to each of the intelligent sensors, i.e. at the lower of the two tiers. The communication capabilities, on the other hand, are utilized to consolidate data and adapted estimators from all the sensors in a cluster. The cluster head is responsible for this consolidation, and subsequently, for the data reconciliation.

This paper will propose and describe in detail the synergy between the two-level partially adaptive data reconciliation and the intelligent sensor network as described above. The robust estimator based on the generalized-t (GT) distribution is used in the data reconciliation. As the name suggest, GT is a general family of distributions covering many commonly encountered distributions such as Gaussian, uniform,

Laplacian and Cauchy. The specific distribution that the GT assumes at a particular instance is determined by the values of GT parameters. Due to this flexibility, in addition to its robustness, the GT distribution is an excellent candidate for the partially adaptive data reconciliation.

A case study involving a subsection of a pilot plant facility will be conducted. Experiment results will demonstrate the feasibility and efficiency of the proposed distributed data reconciliation within the intelligent sensor network framework.

Reference: [1] Y.Y. Joe, Z.Q. Ding, K.V. Ling, J.A. Romagnoli. A Federated Sensor Network Architecture for Data Rectification and Process Monitoring. Submitted to AIChE Annual Meeting, 2005.